



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 96101

TO: Patricia Duffy
Location: cm1/8d05/8e12
Art Unit: 1645
Thursday, June 19, 2003
Case Serial Number: 674277

From: Paul Schulwitz
Location: Biotech-Chem Library
CM1-6B06
Phone: 305-1954

paul.schulwitz@uspto.gov

Search Notes

Examiner Duffy,

See attached results.

If you have any questions about this search feel free to contact me at any time.

Thank you for using STIC search services!

Paul Schulwitz
Technical Information Specialist
STIC Biotech/Chem Library
(703)305-1954

THIS PAGE BLANK (USPTO)

96101

From: Duffy, Patricia
Sent: Sunday, June 08, 2003 10:49 AM
To: STIC-Biotech/ChemLib
Subject: Sequence search 09/674,277

Importance: High

RECEIVED

JUN - 9 2003

(STIC)

In re: 09/674,277

Please search SEQ ID NO:1 and 2.

Please include and interference and commercial database search.

Please print out top 50 hits in each category.

Thanks,

Patricia A. Duffy

CM1-8D05

AU 1645

703-305-7555

Searcher: _____
Phone: _____
Location: _____
Date Picked Up: 6/10
Date Completed: 6/19
Searcher Prep/Review: _____
Clerical: _____
Online time: _____

TYPE OF SEARCH:

NA Sequences: _____
AA Sequences: _____
Structures: _____
Bibliographic: _____
Litigation: _____
Full text: _____
Patent Family: _____
Other: _____

VENDOR/COST (where applic.)

STN: _____
DIALOG: _____
Questel/Orbit: _____
DRLink: _____
Lexis/Nexis: _____
Sequence Sys.: _____
WWW/Internet: _____
Other (specify): _____

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 17, 2003, 15:59:59 ; Search time 2128.45 Seconds
(without alignments)
16148.066 Million cell updates/sec

Title: US-09-674-277-2

Perfect score: 1181

Sequence: 1 ctgcaggagatgagaaaaa.....ttttactttttctctgcag 1181

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 150 summaries

Database :

GenEmbl.:

- 1: gb_ba.:
- 2: gb_hgt.:
- 3: gb_in.:
- 4: gb_cm.:
- 5: gb_ov.:
- 6: gb_pat.:
- 7: gb_ph.:
- 8: gb_pi.:
- 9: gb_pr.:
- 10: gb_ro.:
- 11: gb_sts.:
- 12: gb_sy.:
- 13: gb_un.:
- 14: gb_vi.:
- 15: em_ba.:
- 16: em_fun.:
- 17: em_hum.:
- 18: em_in.:
- 19: em_mu.:
- 20: em_om.:
- 21: em_or.:
- 22: em_ov.:
- 23: em_pat.:
- 24: em_ph.:
- 25: em_pl.:
- 26: em_ro.:
- 27: em_sts.:
- 28: em_un.:
- 29: em_vi.:
- 30: em_hgt_hum.:
- 31: em_hgt_inv.:
- 32: em_hgt_other.:
- 33: em_hgt_mus.:
- 34: em_hgt_pln.:
- 35: em_hgt_rod.:
- 36: em_hgt_mam.:
- 37: em_hgt_vir.:
- 38: em_sy.:
- 39: em_hgtgo_hum.:
- 40: em_hgtgo_mus.:
- 41: em_hgtgo_other.:

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1181	100.0	1181	6	AX011298	Sequence
c 2	1166.2	98.7	7395	1	ECY4PO157	Y11275 E.coli 7.4
c 3	1166.2	98.7	7395	6	AX191728	Sequence
c 4	1166.2	98.7	92077	1	AF074613	AF074613 Escherich
c 5	1166.2	98.7	92077	6	AX191727	Sequence
c 6	1166.2	98.7	92721	1	AB011549	AB011549 Escherich
c 7	1166.2	98.7	92721	6	AX191725	Sequence
c 8	1163	98.5	5612	1	AF043470	Escherich
c 9	328.6	27.8	213494	1	SFPWR100	AL391753 Shigella
c 10	328.6	27.8	221618	1	AF386526	Shigella
c 11	328.6	27.8	221851	1	AF348706	Shigella
c 12	246.6	20.9	3500	1	AF134403	Shigella
c 13	159.4	13.5	1642	1	SHFVIRK	D11025 Shigella fl
c 14	159.4	13.5	1642	6	E06635	E06635 Shigella vi
c 15	154.8	13.1	66714	1	AF326777	AF326777 Shigella
c 16	64.2	5.4	6540	1	BACJOJC	L38424 Bacillus su
c 17	64.2	5.4	10801	6	AX100832	Sequence
c 18	64.2	5.4	23775	1	BACYP1A	L47709 Bacillus su
c 19	64.2	5.4	213680	1	BSUB0012	299115 Bacillus su
c 20	59.2	5.0	13334	1	AE003927	AE003927 Xylella f
c 21	57.2	4.8	14243	1	AE007070	AE007070 Mycobacte
c 22	57.2	4.8	23400	1	MTV021	AL021957 Mycobacte
c 23	57	4.8	14734	1	AF498417	AF498417 Pseudomon
c 24	57	4.8	16736	1	AF035937	AF035937 Pseudomon
c 25	54.8	4.6	237221	1	AP003599	AP003599 Nostoc sp
c 26	52.2	4.4	10197	1	AE008327	AE008327 Agrobacte
c 27	52.2	4.4	12006	1	AE009285	AE009285 Agrobacte
c 28	50.6	4.3	20693	1	SMU52844	-U52844 Serratia ma
c 29	49.6	4.2	8594	1	AE013034	AE013034 Thermoana
c 30	49.6	4.2	199450	9	AC005674	AC005674 Homo sapi
c 31	49	4.1	1135	6	AX433280	Sequence
c 32	48.6	4.1	783	6	AX437703	Sequence
c 33	48.2	4.1	253505	1	AP000007	AP000007 Pyrococcu
c 34	48	4.0	35576	1	SC2D46	AL391406 Streptomy
c 35	47.4	4.0	9297	2	AC092541	AC092541 Homo sapi
c 36	47.4	4.0	10991	1	AE010283	AE010283 Pyrococcu
c 37	47.2	4.0	10148	1	AE012967	AE012967 Chloroblu
c 38	46.2	3.9	783	6	AX437696	Sequence
c 39	46.2	3.9	11884	1	AE004745	AE004745 Pseudomon
c 40	46	3.9	11142	1	AE000805	AE000805 Methanoba
c 41	46	3.9	13401	1	AF498404	AF498404 Pseudomon
c 42	45.8	3.9	300600	1	AP005369	AP005369 Thermosyn
c 43	45.4	3.8	10940	1	AE000469	AE000469 Escherich
c 44	45.4	3.8	110000	2	LMFLCHR32_19	Continuation (20 o
c 45	45.4	3.8	176195	1	ECOW89	U00006 E. coli chr
c 46	45.4	3.8	196146	2	AC023609	AC023609 Mus muscu
c 47	45.4	3.8	254993	2	AC078885	AC078885 Mus muscu
c 48	44.8	3.8	125020	9	AF429315	AF429315 Homo sapi
c 49	44.4	3.8	13413	1	AF498405	AF498405 Pseudomon
c 50	44	3.7	12312	1	AE005627	AE005627 Escherich
c 51	44	3.7	318703	1	AP002567	AP002567 Escherich
c 52	43.4	3.7	5280	1	AF105116	AF105116 Streptoco
c 53	43	3.6	10021	1	AE001736	AE001736 Thermotog
c 54	43	3.6	21409	1	AE004175	AE004175 Vibrio ch
c 55	42.8	3.6	36548	1	MSG81551CS	L78813 Mycobacteri
c 56	42.8	3.6	36548	1	MSG81554CS	L78814 Mycobacteri
c 57	42.8	3.6	38859	1	MLCB268	AL022602 Mycobacte
c 58	42.8	3.6	348450	1	MLEPRTN4	AL583920 Mycobacte
c 59	42.6	3.6	16349	1	AF498420	AF498420 Pseudomon
c 60	42.6	3.6	299350	1	AP005370	AP005370 Thermosyn
c 61	42.4	3.6	2669	1	AF112861	AF112861 Escherich
c 62	42.2	3.6	20385	1	AE008892	AE008892 Salmonell
c 63	42.2	3.6	264050	1	AL627279	AL627279 Salmonell
c 64	41.6	3.5	36734	1	SC6G10	AL049497 Streptomy
c 65	41.6	3.5	1227	6	AX433356	AX433356 Sequence

66	41.6	3.5	13952	1	AF019375	Escherich	AF019375	Escherich	1	AX011298	1181 bp	DNA	linear	PAT 06-SEP-2000
67	41.4	3.5	17453	1	AF498400	Pseudomon	AF498400	Pseudomon	1	AX073342	Sequence			
68	41.4	3.5	341880	1	AF003589	Nostoc sp	AF003589	Nostoc sp	1	AX073342	Sequence			
69	40.8	3.5	9390	6	D87026	Bacillus st	D87026	Bacillus st	6	AE012138	Xanthomon			
70	40.6	3.4	14245	6	AX009415	Sequence	AX009415	Sequence	6	AB033991	Bacillus			
71	40.6	3.4	129908	6	AX059140	Sequence	AX059140	Sequence	6	AC099042	Oryza sat			
72	40.6	3.4	315000	1	RNE603644	Rhizobium	AL603644	Rhizobium	1	AC099042	Oryza sat			
73	40.4	3.4	10029	1	AE013547	Methanob	AE013547	Methanob	1	AC099042	Oryza sat			
74	40.2	3.4	1894	1	SC0243803	Streptomy	AE013547	Streptomy	1	AP003001	Rattus no			
75	40.2	3.4	34983	1	SCM11	Streptomy	AE013547	Streptomy	1	AP003001	Rattus no			
76	40	3.4	24329	1	KPCPS	Klebsiella	D21242	Klebsiella	1	AX026771	Homo sapi			
77	39.8	3.4	14704	1	AE001970	Deinococc	AE001970	Deinococc	1	AX164174	Sequence			
78	39.8	3.4	15670	1	AE006951	Mycobacte	AE006951	Mycobacte	1	AX193704	Sequence			
79	39.8	3.4	37218	1	MTIC2069	Mycobacte	AE006951	Mycobacte	1	SC265	Streptomy			
80	39.8	3.4	158781	2	AC024978	Homo sapi	AC024978	Homo sapi	2	SC265	Streptomy			
81	39.8	3.4	162054	2	AC046201	Homo sapi	AC046201	Homo sapi	2	SC265	Streptomy			
82	39.8	3.4	192974	9	AC087565	Homo sapi	AC087565	Homo sapi	9	SC265	Streptomy			
83	39.8	3.4	202726	9	AC092128	Homo sapi	AC092128	Homo sapi	9	SC265	Streptomy			
84	39.6	3.4	2203	8	NCU46088	Neurospora	U46088	Neurospora	8	SC265	Streptomy			
85	39.6	3.4	2973	8	AF088906	Neurospor	AF088906	Neurospor	8	SC265	Streptomy			
86	39.6	3.4	215889	2	AC092196	Canis fam	AC092196	Canis fam	2	SC265	Streptomy			
87	39.6	3.4	265118	1	CNSPAX06	Pyrococcu	AJ248288	Pyrococcu	1	SC265	Streptomy			
88	39.6	3.4	265118	6	AE041922	Sequence	AE041922	Sequence	6	SC265	Streptomy			
89	39.4	3.3	10206	1	AE013000	Thermoana	AE013000	Thermoana	1	SC265	Streptomy			
90	39.4	3.3	10311	1	AE007621	Clostridi	AE007621	Clostridi	1	SC265	Streptomy			
91	39.4	3.3	10421	1	AE004975	Halobacte	AE004975	Halobacte	1	SC265	Streptomy			
92	38.2	3.3	17013	1	EAAMASL	E.amylovar	X77921	E.amylovar	1	SC265	Streptomy			
93	39.2	3.3	304290	1	CNSPAX05	Pyrococcu	AJ248287	Pyrococcu	1	SC265	Streptomy			
94	39.2	3.3	349980	6	AE041921	Sequence	AE041921	Sequence	6	SC265	Streptomy			
95	39	3.3	11856	1	AE012829	Chlorobiu	AE012829	Chlorobiu	1	SC265	Streptomy			
96	38.8	3.3	10035	1	AE001999	Deinococc	AE001999	Deinococc	1	SC265	Streptomy			
97	38.8	3.3	11172	1	AE010928	Methanoc	AE010928	Methanoc	1	SC265	Streptomy			
98	38.8	3.3	11483	1	AE009930	Pyrocacul	D90913	Synechocyst	1	SC265	Streptomy			
99	38.6	3.3	143308	1	D90913	Synechocyst	D90913	Synechocyst	1	SC265	Streptomy			
100	38.4	3.3	68559	2	AC031981	Homo sapi	AL732467	Mus muscu	2	SC265	Streptomy			
101	38.2	3.2	231258	2	AL732467	Mus muscu	AL732467	Mus muscu	2	SC265	Streptomy			
102	38	3.2	34496	1	SCF62	Streptomy	AL121855	Streptomy	1	SC265	Streptomy			
103	37.8	3.2	57818	2	AC017287	Drosophil	AC017287	Drosophil	2	SC265	Streptomy			
104	37.8	3.2	111448	3	AC005334	Drosophil	AC005334	Drosophil	3	SC265	Streptomy			
105	37.8	3.2	166626	3	AC008318	Drosophil	AC008318	Drosophil	3	SC265	Streptomy			
106	37.8	3.2	301639	9	AE003590	Drosophil	AE003590	Drosophil	9	SC265	Streptomy			
107	37.6	3.2	167405	8	AP002483	Oryza sat	AP002483	Oryza sat	8	SC265	Streptomy			
108	37.6	3.2	175565	8	AP003311	Oryza sat	AP003311	Oryza sat	8	SC265	Streptomy			
109	37.4	3.2	33547	8	PCW0365	Pneumocys	AL592382	Pneumocys	8	SC265	Streptomy			
110	37.2	3.1	1503	8	AF045961	Arabidops	AF045961	Arabidops	8	SC265	Streptomy			
111	37.2	3.1	1533	8	AF045961	Arabidops	AF045961	Arabidops	8	SC265	Streptomy			
112	37.2	3.1	7581	1	AB005901	Streptomy	AB005901	Streptomy	1	SC265	Streptomy			
113	37.2	3.1	125020	9	AF429315	Homo sapi	AF429315	Homo sapi	9	SC265	Streptomy			
114	37.2	3.1	300950	6	AX433315	Sequence	AX433315	Sequence	6	SC265	Streptomy			
115	37	3.1	1108	6	AX433315	Sequence	AX433315	Sequence	6	SC265	Streptomy			
116	37	3.1	17575	1	AE004913	Pseudomon	AE004913	Pseudomon	1	SC265	Streptomy			
117	37	3.1	192657	1	AP005377	Thermosyn	AP005377	Thermosyn	1	SC265	Streptomy			
118	37	3.1	199317	2	AC100400	Mus muscu	AC100400	Mus muscu	2	SC265	Streptomy			
119	36.8	3.1	18848	4	CCY14852	Oryctolagu	Y14852	Oryctolagu	4	SC265	Streptomy			
120	36.8	3.1	40352	1	SCC61A	Streptomy	AL356595	Streptomy	1	SC265	Streptomy			
121	36.8	3.1	118507	8	ATF7J8	Arabidops	AP001512	Arabidops	8	SC265	Streptomy			
122	36.8	3.1	300550	1	AP001512	Bacillus	AP001512	Bacillus	1	SC265	Streptomy			
123	36.6	3.1	654	6	AX280161	Sequence	AX280161	Sequence	6	SC265	Streptomy			
124	36.6	3.1	1372	1	AF061562	Mycobacte	AF061562	Mycobacte	1	SC265	Streptomy			
125	36.6	3.1	11119	1	AE013060	Thermoana	AE013060	Thermoana	1	SC265	Streptomy			
126	36.6	3.1	14725	1	AE006956	Mycobacte	AE006956	Mycobacte	1	SC265	Streptomy			
127	36.6	3.1	40051	1	MSGV224	Mycobacte	AD000004	Mycobacte	1	SC265	Streptomy			
128	36.6	3.1	40838	1	MTY25D10	Mycobacte	Z95558	Mycobacte	1	SC265	Streptomy			
129	36.4	3.1	175127	2	AC121333	Homo sapi	AC121333	Homo sapi	2	SC265	Streptomy			
130	36.4	3.1	183994	2	AC069331	Homo sapi	AC069331	Homo sapi	2	SC265	Streptomy			
131	36.2	3.1	741	6	AX280220	Sequence	AX280220	Sequence	6	SC265	Streptomy			
132	36.2	3.1	1704	6	AX193681	Sequence	AX193681	Sequence	6	SC265	Streptomy			
133	36.2	3.1	10732	6	E32986	Gene encodi	E32986	Gene encodi	6	SC265	Streptomy			
134	36.2	3.1	67521	2	AC103845	Homo sapi	AC103845	Homo sapi	2	SC265	Streptomy			
135	36.2	3.1	150052	9	AC018525	Homo sapi	AC018525	Homo sapi	9	SC265	Streptomy			
136	36.2	3.1	178736	9	AC103853	Homo sapi	AC103853	Homo sapi	9	SC265	Streptomy			
137	36.2	3.1	181688	2	AP005354	Homo sapi	AP005354	Homo sapi	2	SC265	Streptomy			
138	36.2	3.1	192581	2	AC121044	Rattus no	AC121044	Rattus no	2	SC265	Streptomy			

ALIGNMENTS

RESULT 1
AX011298

LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE

ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL

FEATURES
Source

BASE COUNT
ORIGIN

Query Match
Best Local Similarity
Matches 1181; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGCAAGGAGATGCAAAAAGCAAAAATTTCCCATCCCGCGCTCCAGCTG 60

Db 1 CTGCAAGGAGATGCAAAAAGCAAAAATTTCCCATCCCGCGCTCCAGCTG 60

QY 61 AAGTAGGCGCTTCTGTCGGGTATTTAAATGCAATGACCGTCCCGCTATTTAAACAATG 120

Db 61 AAGTAGGCGCTTCTGTCGGGTATTTAAATGCAATGACCGTCCCGCTATTTAAACAATG 120

QY 121 TGAATAATACCTCCGTTACCGGAAACCGTGAACAAAATTCGGGCTGAAAGAGGATCC 180

Db 121 TGAATAATACCTCCGTTACCGGAAACCGTGAACAAAATTCGGGCTGAAAGAGGATCC 180

QY 181 GCGGTATCTGTGTGATTTCCCTTAGCTGACCTAGCCAGACAGACAAATGATCTGTGCGG 240

Db 181 GCGGTATCTGTGTGATTTCCCTTAGCTGACCTAGCCAGACAGACAAATGATCTGTGCGG 240

QY 241 TTCTGTTAATATCAACCGGTACTCAATATCTTCTGTCGGCTGGCTGCCATCATCCGGA 300

Db 241 TTCTGTTAATATCAACCGGTACTCAATATCTTCTGTCGGCTGGCTGCCATCATCCGGA 300

QY 301 ACGGTCGCGTGGGATATAAATCGGCAGTCGGCGGTCCATGCAGACACATCCGCCA 360

Db 301 ACGGTCGCGTGGGATATAAATCGGCAGTCGGCGGTCCATGCAGACACATCCGCCA 360

QY 361 CGGGTAACAGCGTCCCTGTCACATTTCTTCTGAATGACATCAGGATCCCGCCCTCTCAC 420

Db 361 CGGGTAACAGCGTCCCTGTCACATTTCTTCTGAATGACATCAGGATCCCGCCCTCTCAC 420

QY	421	TGCGGATACGGGACGCGGAGACGATGACGCTTACGCGATACCATACCAACGCTTCAT	480
Db	421		
QY	481	TTTCCGAAGGATGACACACACTGGCAATCGGTAGACCGGTACGCTGGGAAAGG	540
Db	481		
QY	541	CACCTGCCATTAAACACATCCGCTCATTTCCAGAGTGTCTGCTGACGACGAGTG	600
Db	541		
QY	601	CTTCGTTATTCACGCGCGCGCGCCACACGAGCAGCGAAATGATTTCCCTTCATCT	660
Db	601		
QY	661	TCAGCTGATACATACACGACGATAAATTCATGCTCTTTTCGGGACGTAGCATCCCCA	720
Db	661		
QY	721	CCTGAACGATAAGCGGAACATTGTCTGCTGATGACGCCAGCGGTGATGACGGGTA	780
Db	721		
QY	781	ACGGTCCGATGCTTCATTATGCAATGCGGCGCAGTCGAACCCGGTGAATACCGTTA	840
Db	781		
QY	841	CGGCTGCTGACACCTTCGCGCATCAGATGCGCCATCATGGTGAGATAGCACACAA	900
Db	841		
QY	901	TGAATCACACAGATAATTCAGGAAACGTTCTGCTTACGGGTGATGATAGTTTTT	960
Db	901		
QY	961	GTCTGACATAGTGAACGGTGACACATATCAGACGCTCAGTCTGCTATATTAATCT	1020
Db	961		
QY	1021	CATGGCCACTATGGCAGATGACCATCATAGGTTAAATCCCGATAATCCCGTGAAGTC	1080
Db	1021		
QY	1081	TGAGGATGAAGGAGGTGAAGCTCTTCTGAAAGGAATAAAGTGACATCATGCCCTC	1140
Db	1081		
QY	1141	TTTTTCTGCTCCGAGCAATTTTACTTTTTTCTCTGCAG	1181
Db	1141		
RESULT 2	EC74P0157/c	7395 bp	DNA
LOCUS	EC74P0157/c	7395 bp	DNA
DEFINITION	E.coli 7.4 kb DNA from plasmid p0157.		
ACCESSION	Y11275		
VERSION	Y11275.1		
KEYWORDS	msbB gene; unidentified protein.		
SOURCE	Escherichia coli.		
ORGANISM	Escherichia coli.		
REFERENCE	Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;		
AUTHORS	1 (bases 1 to 7395)		
JOURNAL	Bruder, W.		
AUTHORS	Unpublished		
TITLE	2 (bases 1 to 7395)		
JOURNAL	Bruder, W.		
	Submitted (14-FEB-1997) W. Bruder, Universitaet Wuerzburg,		
	Institut fuer Hygiene und Mikrobiologie, Josef-Schneider-Strasse 2,		

D- 97080	Wuerzburg, FRG
Related	sequences X97542, M77039, X02311.
Location/Qualifiers	
1. .7395	
/organism="Escherichia coli"	
/strain="EDL933"	
/serotype="O157:H7"	
/db_xref="taxon:562"	
/clone="psm7"	
/plasmid="p0157"	
774. .1595	
/codon_start=1	
/transl_table=11	
/product="hypothetical protein"	
/protein_id="CAA72138.1"	
/db_xref="GI:4127813"	
/translation="MLNARHLPLVLMYHVSRCPLVLTSPETFCFCEOMKLADNNWRTV	
TSALEYFYGGTLPKSYMLTFDGYLDNWLRAWPVLYEYLNHAFILFTLGLIGGN	
VSRQEQEYSHRDCERLIAADNRSEVREMRDGLVEFVHTTHSHKRWDRLS	
VSRQEQRLMKEDILVGKQCLEKGLGFCSCWPEGYNYRDYINLAGLIGFSYLT	
ERRMCPENGSLRIGRISTKEREHSGWLKRRFLYTTPLFSSVLAHLHKGPRLPDN"	
1595. .2701	
/codon_start=1	
/transl_table=11	
/product="hypothetical protein"	
/protein_id="CAA72139.1"	
/db_xref="GI:4127814"	
/translation="MMKILFTSSSDIGGOELQALQMTALOKOGHSLVLLACREKSKI	
APEARRGHDVTFIPPRNSLHPLSILRLRIIGFEPDLVICHSGHDSNLAGLSRLIC	
CHRFISVROKTYITRTKTRFSLNLCDFIVPSSAMMAHLMAEGVTPVTVIIPGDDW	
PAHNEAMRPLHIIHAAASADNVPLIVGVMLRPEKHEFMRLVYQUMKMGKFSR	
WLVVGAREYEARLQOQTHSGMLGALFALPFIYRIASVYVAMPSENEAFGM	
VLAASDVGPVVIASITGGTIPVIOKNVTCTLLPVGDSVAWTGALRDFLSRPERFRM	
AASARDIEYRFDINRTAQIIVSLASQAKCNR"	
2791. .4512	
/codon_start=1	
/transl_table=11	
/product="hypothetical protein"	
/protein_id="CAA72140.1"	
/db_xref="GI:4127815"	
/translation="MHLNTGONRPTFSWALGWAIFYFGFFSTLLOVIFSSGSGTN	
GIRDSLLFSLMLIPVLYPDRIKIIAAVVGFTLWGTSLAACLYELYGHEFSQVLF	
VMEFTNARAGEYFSQYFSLKLLLSLYTAVSVFLWTLRLPVYIPLPWRRIYSVFL	
YALLHPVLKLSIROEPLNDTLGLASRMPEAPWQFVSSYYHQOALNLTFLNE	
NSALPLGLNRDESQERPTLVIGESTQREMSLYGLYRETTPDLARKTDPLGT	
VNNVVASRPTTEALQALTFANEKNPDLYTPQSLMMKQAGKFTWITNQOIT	
ARNTMLTVESROTDROYVMNQRTQSAVEDNLVLPFREVLDNPAKKLIIVHLLGT	
HKYKYVPDGGGREDGITHIPTGLNAKELEVYNDYDNANVFNHVASLINDFRAT	
APDGFLLYSDHGEVEYDPPYKTQGRNEDNTPRPMYVFFLLWTSEKWHAAHPDFS	
QYVDRKYSIAELIHTWSDLAGLTYDGTPTSLVNPQFRETTRWIGNPYKKNGLTDFD	
TLPYGEPT	
4586. .5584	
/gene="msbB"	
4586. .5584	
/gene="msbB"	
/codon_start=1	
/transl_table=11	
/product="msbB protein"	
/protein_id="CAA72141.1"	
/db_xref="GI:4127816"	
/translation="WYPPSAVLCYHNEISRQIPVNMKNINTEFIPRNLTLCPPTWM	
TWTGIGIICVAMVPPALRDPLGLKGLMVLGRKSGARALINLSLCPPEYSDKEKE	
NIVDAMFATASMAVLMALALSGDPKISHRIWNGLEIVKMAQKNKQYIFLVPVHW	
GVDIIPAMLAQSGKMAAFHMQRPVYVWNSVRRFGCKLHARNNDGIASVRSVR	
QGYWGYLPDQDHGPEFSFADPFATYKATLPVIGRLSRISGARIIPLPVYDGKTHH	
LTHVSPPLAIRQKSDAHARQINEVENVFVRPHEQYTWILKLLATRKEGEDPY"	
5645. .5820	
/note="IS3-like sequence"	
6465. .57395	
/codon_start=1	
/transl_table=11	
/product="hypothetical protein"	
/protein_id="CAA72142.1"	

misc_feature

CDS

```
/db_xref="GI:4127817"
/translation="MNTKMERWRTPMKLYLSCTLAPLAIGVFSATAADNNSAIYF
NTSOPINDLAGSAAEVKFAQSILPAHPKREGDSQPHL7SLRSLKLLVPRVADRTTP
VOYARDNNKILGTLTYPPSLPDTIYHLGDVPEGGIDFTPHNGTKKILNTVAENV
KLSDAGSSSHSLHNNALVEIHTANGRWVDIYLPQDPLEGKMWRFVSSAGYSSTV
FYGRKVTLSVGNLTFLKYVNGWFRSGELENNRITYAQHWSAELPAHWIVPLNLV
IKQNLSEGRUNDIKIGAPGBELLHTIDIGMLTTP"
BASE COUNT 1929 a 1575 c 1815 g 2076 t
ORIGIN

Query Match 98.7%; Score 1166.2; DB 1; Length 7395;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1168; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 11 TGGAAAAAGCCAAATAAAAAATGCCATCCAGCGCGCTCCAGCTGAAAGTAGGCC 70
DB 2875 TGGAAAAAGCCAAATAAAAAATGCCATCCAGCGCGCTCCAGCTGAAAGTAGGCC 2816
QY 71 TGTTCGTGCGGTATTTAAATGCAATGACCGTCCCGTATTTAAACAATGTGATAATTA 130
DB 2815 TGTTCGTGCGGTATTTAAATGCAATGACCGTCCCGTATTTAAACAATGTGATAATTA 2756
QY 131 CTCGGTACCGGAAAACCGCTGAAACAAATTCGGGCTGAAAGAGGATCCGCGTTATCT 190
DB 2755 CTCGGTACCGGAAAACCGCTGAAACAAATTCGGGCTGAAAGAGGATCCGCGTTATCT 2696
QY 191 GTTGCAATTTCCCTTAGCCTGACTAGCCAGACACAATGATGTGCGGTTCTGTTAAT 250
DB 2695 GTTGCAATTTCCCTTAGCCTGACTAGCCAGACACAATGATGTGCGGTTCTGTTAAT 2636
QY 251 ATCAAAACCGGTACTCAATATCTTCTGCGCTGGCTGCATCATCCGGAAGGTTCCGG 310
DB 2635 ATCAAAACCGGTACTCAATATCTTCTGCGCTGGCTGCATCATCCGGAAGGTTCCGG 2576
QY 311 TCGGGATAAAAAATCGCGAGTGC CGCGGTTCATGAGACACATCCCGGAGGTTAACAG 370
DB 2575 TCGGGATAAAAAATCGCGAGTGC CGCGGTTCATGAGACACATCCCGGAGGTTAACAG 2516
QY 371 CGTCCCTGTACATCTTCTGAATGACATCAGGAGTCCGCGCTCTCACTGCGGATPAC 430
DB 2515 CGTCCCTGTACATCTTCTGAATGACATCAGGAGTCCGCGCTCTCACTGCGGATPAC 2456
QY 431 GGGCAGCGCGGAGACTGACGCTTCAGCAGTACCATACCAACGCTTCATTTCCGAGG 490
DB 2455 GGGCAGCGCGGAGACTGACGCTTCAGCAGTACCATACCAACGCTTCATTTCCGAGG 2396
QY 491 CATGACCACCACTGGCAATCCGGTAGACCGGTAGCGGTGAGGAAAGGCACTGCGCAT 550
DB 2395 CATGACCACCACTGGCAATCCGGTAGACCGGTAGCGGTGAGGAAAGGCACTGCGCAT 2336
QY 551 TAACACATCTCCGTCTATCCAGGTGTTCTGTCTGTGACGACAGAGTGTCTGTTATTC 610
DB 2335 TAACACATCTCCGTCTATCCAGGTGTTCTGTCTGTGACGACAGAGTGTCTGTTATTC 2276
QY 611 TTCAGCGCGCGCGCCACCGAGCGAGCGAATGATTCCTTCCATCTTCAGCTGATA 670
DB 2275 TTCAGCGCGCGCGCCACCGAGCGAGCGAATGATTCCTTCCATCTTCAGCTGATA 2216
QY 671 CAATACAGCGAGCAATAATTCCTTTTCGGGAGCTAGCATCCCGACCTGAACGAT 730
DB 2215 CAATACAGCGAGCAATAATTCCTTTTCGGGAGCTAGCATCCCGACCTGAACGAT 2156
QY 731 AAGCGGAACATGTGCTGTATGACGCGCGAGCGGTGGATATGACAGGGGTAAACGTCGCAT 790
DB 2155 AAGCGGAACATGTGCTGTATGACGCGCGAGCGGTGGATATGACAGGGGTAAACGTCGCAT 2096
QY 791 GCGTTCATATGCAATCGCGCCAGTCCGAACCGGTGGATATACCGTACCCTGTCCT 850
DB 2095 GCGTTCATATGCAATCGCGCCAGTCCGAACCGGTGGATATACCGTACCCTGTCCT 2036
QY 851 GACACCTTCGCGCATACAGTGGCCATCATGGGTGAGATAGGACCAACAATGAATCACA 910
DB 2035 GACACCTTCGCGCATACAGTGGCCATCATGGGTGAGTAGGACCAACAATGAATCACA 1976
```

```
QY 911 CAGATAAATTCAGGAAAAAGCTTCTGGTCTTACGGGTGATGTAGGTTTTTGTCTGACAAT 970
DB 1975 CAGATAAATTCAGGAAAAAGCTTCTGGTCTTACGGGTGATGTAGGTTTTTGTCTGACAAT 1916
QY 971 AGTGAACCGGTGACAGCATATCAGACGGCTCAGTCTGCTATATTTACTGTCTAGCCACT 1030
DB 1915 ACTGAACCGGTGACAGCATATCAGACGGCTCAGTCTGCTATATTTACTGTCTAGCCACT 1856
QY 1031 ATGGCAGATGACAGCATCAGGTTTAAATTCGCCGATAATCCCGTCAAGTCTCAGGATGA 1090
DB 1855 ATGGCAGATGACAGCATCAGGTTTAAATTCGCCGATAATCCCGTCAAGTCTCAGGATGA 1796
QY 1091 AGGAAGGTGAAGGCTTCTCTGAAAGGAATAAAGTGAACATCATGCCCTCTTTTCTGGC 1150
DB 1795 AGGAAGGTGAAGGCTTCTCTGAAAGGAATAAAGTGAACATCATGCCCTCTTTTCTGGC 1736
QY 1151 TTCCCGGAGCAATTTTACTTTTTTCTCTGCAG 1181
DB 1735 TTCCCGGAGCAATTTTACTTTTTTCTCTGCAG 1705

RESULT 3
AX191728/c
LOCUS AX191728 7395 bp DNA linear PAT 15-AUG-2001
DEFINITION Sequence 10 from Patent WO0149775.
ACCESSION AX191728
VERSION AX191728.1 GI:15209897
KEYWORDS
SOURCE
ORGANISM Escherichia coli.
Bacterias; Proteobacteria; gamma subdivision; Enterobacteriaceae;
Escherichia.
REFERENCE 1 (bases 1 to 7395)
AUTHORS Iversen,P.L.
TITLE Antisense antibacterial cell division composition and method
JOURNAL Patent: WO 0149775-A 10 12-JUL-2001;
AVI Biopharma, Inc. (US)
FEATURES
source
Location/Qualifiers
/organism="Escherichia coli"
/db_xref="taxon:562"
BASE COUNT 1929 a 1575 c 1815 g 2076 t
ORIGIN

Query Match 98.7%; Score 1166.2; DB 6; Length 7395;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1168; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 11 TGGAAAAAGCCAAATAAAAAATTCGCCATCCAGCGCGCTCCAGCTGAAAGTAGGCC 70
DB 2875 TGGAAAAAGCCAAATAAAAAATTCGCCATCCAGCGCGCTCCAGCTGAAAGTAGGCC 2816
QY 71 TGTTCGTGCGGTATTTAAATGCAATGACCGTCCCGTATTTAAACAATGTGATAATTA 130
DB 2815 TGTTCGTGCGGTATTTAAATGCAATGACCGTCCCGTATTTAAACAATGTGATAATTA 2756
QY 131 CTCGGTACCGGAAAACCGCTGAAACAAATTCGGGCTGAAAGAGGATCCGCGTTATCT 190
DB 2755 CTCGGTACCGGAAAACCGCTGAAACAAATTCGGGCTGAAAGAGGATCCGCGTTATCT 2696
QY 191 GTTGCAATTTCCCTTAGCCTGACTAGCCAGACACAATGATGTGCGGTTCTGTTAAT 250
DB 2695 GTTGCAATTTCCCTTAGCCTGACTAGCCAGACACAATGATGTGCGGTTCTGTTAAT 2636
QY 251 ATCAAAACCGGTACTCAATATCTTCTGCGCTGGCTGCATCATCCGGAAGGTTCCGG 310
DB 2635 ATCAAAACCGGTACTCAATATCTTCTGCGCTGGCTGCATCATCCGGAAGGTTCCGG 2576
QY 311 TCGGGATAAAAAATCGCGAGTGC CGCGGTTCATGAGACACATCCCGGAGGTTAACAG 370
DB 2575 TCGGGATAAAAAATCGCGAGTGC CGCGGTTCATGAGACACATCCCGGAGGTTAACAG 2516
```

```
QY 371 CGTCCCTGTCACATTTCTTCTGATGACATCAGGGATCCGCCGCTCTACTGGCGATAAC 430
Db 2515 CGTCCCTGTCACATTTCTTGAATGACATCAGGGATCCGCCGCTCTACTGGCGATAAC 2456
QY 431 GGCACGCCGGAGACTGACGCTTCAGCCAGTACCATACCAAGCGCTTCATTTCCGAAGG 490
Db 2455 GGCACGCCGGAGACTGACGCTTCAGCCAGTACCATACCAAGCGCTTCATTTCCGAAGG 2396
QY 491 CATGACACACACATGGCAATCCGCTAGACCGGTAAACGCTGGGAAAGGCGACCTGCCAT 550
Db 2395 CATGACACACACATGGCAATCCGCTAGACCGGTAAACGCTGGGAAAGGCGACCTGCCAT 2336
QY 551 TAACACATCTCCGCTCATTTCCAGGTGTTCTGCTGCTACGACGAGCTGCTTCGTATTC 610
Db 2335 TAACACATCTCCGCTCATTTCCAGGTGTTCTGCTGCTACGACGAGCTGCTTCGTATTC 2276
QY 611 TTCACGCCCGGGCCGCCACACGAGCGAGGAATGATTTCCCTTCCATCTTCAGCTGATA 670
Db 2275 TTCACGCCCGGGCCGCCACACGAGCGAGGAATGATTTCCCTTCCATCTTCAGCTGATA 2216
QY 671 CAATACACGACGATAAATTCATGTCCTTTTTCGGGACGTAGCATCCCCACCTGAACGAT 730
Db 2215 CAATACACGACGATAAATTCATGTCCTTTTTCGGGACGTAGCATCCCCACCTGAACGAT 2156
QY 731 AAGCGGAACATTTGCTGCTGATGACGCCAGCGGTGGATATGACGGGTAAACGGTCCAT 790
Db 2155 AAGCGGAACATTTGCTGCTGATGACGCCAGCGGTGGATATGACGGGTAAACGGTCCAT 2096
QY 791 GGCTTCATTTATGCAATCGCGGCGAGTCGAAACCCGGTGAATTAACGGTTCACGGTCTCT 850
Db 2095 GGCTTCATTTATGCAATCGCGGCGAGTCGAAACCCGGTGAATTAACGGTTCACGGTCTCT 2036
QY 851 GACACCTTCGCCATCAGATGCGCCATCATGCGTGGATAGAGGACACAAATGAATCACA 910
Db 2035 GACACCTTCGCCATCAGATGCGCCATCATGCGTGGATAGAGGACACAAATGAATCACA 1976
QY 911 CAGATAATTCAGGAAAGCTTCTGCTTACGGGTGATGATAGTTTTCCTGACAT 970
Db 1975 CAGATAATTCAGGAAAGCTTCTGCTTACGGGTGATGATAGTTTTCCTGACAT 1916
QY 971 AGTGAAGCGGTGACAGCATATCAGCGCTCAGTCTCTGCTATATTAATCTGTCATGCCACT 1030
Db 1915 AGTGAAGCGGTGACAGCATATCAGCGCTCAGTCTCTGCTATATTAATCTGTCATGCCACT 1856
QY 1031 ATGGCAGATGACAGATCAGTTAAATTCGCCGATAATCCGTCGAAGTCTGAGGATGA 1090
Db 1855 ATGGCAGATGACAGATCAGTTAAATTCGCCGATAATCCGTCGAAGTCTGAGGATGA 1796
QY 1091 AGGAAGGTGAAGCTGTTCTGNAAGGATAAAGTGACATCATGCCCTCTTTTCTGGC 1150
Db 1795 AGGAAGGTGAAGCTGTTCTGNAAGGATAAAGTGACATCATGCCCTCTTTTCTGGC 1736
QY 1151 TTCCGGAGCAATTTTACTTTTCTCTGCGAG 1181
Db 1735 TTCCGGAGCAATTTTACTTTTCTCTGCGAG 1705
```

RESULT 4

```
AF074613/c 92077 bp DNA circular BCT 04-NOV-1998
LOCUS
DEFINITION Escherichia coli O157:H7 plasmid pO157, complete sequence.
ACCESSION AF074613
VERSION AF074613.1 GI:3822114
KEYWORDS
SOURCE Escherichia coli O157:H7.
ORGANISM Escherichia coli O157:H7
Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
REFERENCE 1 (bases 1 to 92077)
AUTHORS Burland,V., Shao,Y., Perna,N.T., Plunkett,G., Sofia,H.J. and
Blattner,F.R.
TITLE The complete DNA sequence and analysis of the large virulence
plasmid of Escherichia coli O157:H7
```

```
JOURNAL Nucleic Acids Res. 26 (18), 4196-4204 (1998)
MEDLINE 98391744
PUBMED 9722640
REFERENCE 2 (bases 1 to 92077)
AUTHORS Burland,V., Shao,Y., Perna,N.T., Plunkett,G. III, Sofia,H.J. and
Blattner,F.R.
TITLE Direct Submission
JOURNAL Genetics, University of Wisconsin, 445
Submitted (25-JUN-1998) Madison, WI 53706, USA
FEATURES
Source
1. 92077
/organism="Escherichia coli O157:H7"
/strain="EDL933"
/serotype="O157:H7"
/db_xref="taxon:83334"
/lab_host="Escherichia coli C600"
/plasmid="pO157"
1. .561
/gene="fno"
/notes="L7001"
1. .561
/gene="fno"
/notes="97 pct identical amino acid sequence and equal
length to F101_ECOLI SW: P22707"
/codon_start=1
/transl_table=11
/product="fertility inhibition protein (conjugal transfer
repressor)"
/protein_id="AAC70069.1"
/db_xref="GI:3822115"
/translation="MAEQRPVLTLRKTEGETPVRSRKTIINVTTPPKVKVKQKLA
EKAAREAEALAAKQAQALSIYLNPTQDEAVNTLKPWPGLDGDTPLRLACGIRD
VILEDAQNIPLSHKRLRAKAITRESYLCAMKAGACRYDTGYVYVTHISQEEBA
YAAARLDKIRQRNRIKAEQLQAVLDER"
653. 681
/notes="predicted sigma 70 promoter; score of 52%"
/gene="L7002"
698. 949
/gene="L7002"
698. 949
/gene="L7002"
/notes="54 pct identical (1 gap) to 23 residues of an
approx. 269 aa protein EXOK_RHIME SW: P33693"
/codon_start=1
/transl_table=11
/product="unknown"
/protein_id="AAC70070.1"
/db_xref="GI:3822116"
/translation="MDEYVHGTVRSVTSVPEAGPLFWKSVSDAGWKRQKHGDLPLV
HGLTGSUPEKGLNTATGAAEGGNEKSLHYSDSEG"
1025. 1052
/notes="predicted sigma 70 promoter; score of 62%"
1151. 1612
/gene="L7003"
1151. 1612
/gene="L7003"
/notes="98 pct identical and equal length to YFI3_ECOLI SW:
Q99342"
/codon_start=1
/transl_table=11
/product="hypothetical protein 15.6 kDa protein in fno 3'
region precursor"
/protein_id="AAC70071.1"
/db_xref="GI:3822117"
/translation="MBKYIPLVLFISWPVLCADIHGRVVRVLDGDTIEVMSRKAIR
IRLVNIDAPEKKQDYGWSTDMKSLVACKTVITYTFQDRDRYGRMLGOVYAPGMVNV
QPMVRAGAAWYVEQYINTDPVLVPLQNEARQQKRLGWSADPVPFWIWRHK"
1390. 1419
/notes="predicted sigma 70 promoter; score of 56%"
1658. 1867
/gene="L7004"
1658. 1867
/gene="L7004"
/notes="55 pct identical (0 gaps) to 66 residues of an
```

```

approx. 72 aa protein HHA_ECOLI SW: P23870"
/codon_start=1
/transl_table=11
/product="putative hemolysin expression modulating
protein"
/protein_id="AAC70072.1"
/db_xref="GI:3822118"
/translation="MEYKQEWLYQLRCCSSVNTLEKIIHKNRDSLSTERESFNAA
DHRAELITGKLYDRIPKEIKVYR"
1899..1927
/feature="predicted sigma 70 promoter; score of 56%"
1905..2243
/feature="L7005"
1905..2243
/feature="L7005"
/feature="98 pct identical (0 gaps) to 57 residues of an
approx. 200 aa protein; plasmid R100 miniplasmid pSM1 ORF
4, TRSW: Q52340"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAC70073.1"
/db_xref="GI:3822119"
/translation="MKLIIFILVLIIAALLIRILRSVNOHSPLLMOLHAAGIRGTD
AERTLSGSEVDASHLERRRRDILLEVLQRGIFLLRSHDARKLLQMTGEWLNTTG
AAQSPHRS"
complement(2240..2327)
/feature="initiation site; Escherichia coli plasmid R100 ss1B
gene; 99 pct identical to ss1B locus R10SS1B accession
D90185"
2416..2421
/feature="100 pct identical (0 gaps) to the -35 region at
23..28 locus ECNR1REP accession X02302"
2439..2445
/feature="100 pct identical (0 gaps) to the -10 region at
46..52 locus ECNR1REP accession X02302"
2473..2478
/feature="100 pct identical (0 gaps) to RBS at 80..85 locus
ECNR1REP accession X02302"
2483..2737
/feature="cpb2"
2483..2737
/feature="L7006"
2483..2737
/feature="cpb2"
/feature="98 pct identical amino acid sequence and equal
length to CPB2_ECOLI SW: P03847"
/codon_start=1
/transl_table=11
/product="CcpB protein (RepA2 protein)"
/protein_id="AAC70074.1"
/db_xref="GI:3822120"
/translation="MSQTENAVTVSSGAKRAYRKGPNPLSDAEKQRLSVARKRPFKEV
KVLEPKYKAMLMQMCHEDGLTQAEVLTKLSEAKRCV"
2754..2782
/feature="predicted sigma 70 promoter; score of 59%"
2973..3047
/feature="L7007"
2973..3047
/feature="L7007"
/feature="91 pct identical (0 gaps) to 24 residues of an
approx. 24 aa protein uORF P30REPFIC, accession M16167,
translationally coupled to replication initiation protein"
/codon_start=1
/transl_table=11
/product="replication initiation protein"
/protein_id="AAC70075.1"
/db_xref="GI:3822121"
/translation="MLGKVDQDFLLCSILLIRIVSAGWCD"
3032..3035
/feature="100 pct identical (0 gaps) to RBS at 641..644 locus
ECNR1REP accession X02302"
3040..3897
/feature="rep2"

```

```

/feature="L7008"
3040..3897
/feature="rep2"
/feature="96 pct identical amino acid sequence and equal
length to REP2_ECOLI SW: P03066"
/codon_start=1
/transl_table=11
/product="replication initiation protein"
/protein_id="AAC70076.1"
/db_xref="GI:3822122"
/translation="MTDIQOTVYROVKPNPVPETREGAGTLKCEKLMKAVGFTSR
FDFAIHAHRSKGLRRMPVLRRAIDALQGLCFHYDPLANRVQCSITLALTECG
LATESAGKLSITRATLTFUSELGLITYQTEYDPLGICYPTDITTFALFALDV
SEDAVAARRSRVWENLRKKQGLDTLGMDELIKAWRFVREFRSOTTELKSGIK
RARARDRAGRQDIVTLVKQLTREISGREFTNAREAVKREVERVERMILSRNRK
YSRLATASP"
3847..4095
/standard_name="CIS"
/feature="88 pct identical to locus ECIS accession X12587,
required for cis-activation of orf1 by the replication
initiation protein"
4072..4080
/feature="dnaA site; 100 pct identical (0 gaps) to locus
ECNR1REP at (1682..1690) accession X02302"
4084..4232
/standard_name="ori"
/feature="89 pct identical to ori (1094..1242); minimum
segment for replication of E. coli IncFII plasmid NR1
ECREPAL X12776"
/direction=right
/promoter
4169..4197
/feature="predicted sigma 70 promoter; score of 56%"
4259..4453
/feature="L7009"
4259..4453
/feature="L7009"
/feature="95 pct identical to (0 gaps) 64 residues of a 128
aa protein REP44 locus ECRC1 accession V00351"
/codon_start=1
/transl_table=11

```

```

Query Match      98.7%; Score 1166.2; DB 1; Length 92077;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1168; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 11 TGGAAAAAAGCCAAAATAAAAAATGCCATCCAGCGCGTCCAGTGAAGTAGGCC 70
DB 19424 TGGAAAAAAGCCAAAATAAAAAATGCCATCCAGCGCGTCCAGTGAAGTAGGCC 19365
QY 71 TCTTCTGTCGGTATTAAATGCAATGACCGTCCCGCTATTAAACAATGTGATAATTA 130
DB 19364 TCTTCTGTCGGTATTAAATGCAATGACCGTCCCGCTATTAAACAATGTGATAATTA 19305
QY 131 CTCGGTTACCGAAAAACCGCTGAACAAAATTCGGGCTGAAAAGAGGATCCCGGTATCT 190
DB 19304 CTCGGTTACCGAAAAACCGCTGAACAAAATTCGGGCTGAAAAGAGGATCCCGGTATCT 19245
QY 191 GTTGCAATTTCCCGTTAGCTAGTCCAGACACAAATGATCTGTGCGCTTCTGTTAAT 250
DB 19244 GTTGCAATTTCCCGTTAGCTAGTCCAGACACAAATGATCTGTGCGCTTCTGTTAAT 19185
QY 251 ATCAACCGGTACTCAATATCTTCTGTGGCGTGGCTGCCATCATCCGGAAGCGTTCGGG 310
DB 19184 ATCAACCGGTACTCAATATCTTCTGTGGCGTGGCTGCCATCATCCGGAAGCGTTCGGG 19125
QY 311 TCGGGATAAAAAATCGCGAGTGGCGGTCATCCATGACACACATCCCCCGGGTAAACAG 370
DB 19124 TCGGGATAAAAAATCGCGAGTGGCGGTCATCCATGACACACATCCCCCGGGTAAACAG 19065
QY 371 CGTCCCTGTACATTTCTTCTGAATGACATCAGGATCCCGCGCTCTCACTGGCGATAAC 430
DB 19064 CGTCCCTGTACATTTCTTCTGAATGACATCAGGATCCCGCGCTCTCACTGGCGATAAC 19005

```

```
QY 431 GGCACGCCGAGACTGACGCTTCAGCAGTACCATACCAACGCTTCATTTCCGAAGG 490
Db 19004 GGCACGCCGAGACTGACGCTTCAGCAGTACCATACCAACGCTTCATTTCCGAAGG 18945
QY 491 CATGACACACACTGGCAATCGGTAGACCGTAAACGCTGGGAAAGGACCTGCGCAT 550
Db 18944 CATGACACACACTGGCAATCGGTAGACCGTAAACGCTGGGAAAGGACCTGCGCAT 18885
QY 551 TAACACATCTCCGCTCATTTCCCAAGGTGTTCTGCTGACGAGACGCTGCTCGATTC 610
Db 18884 TAACACATCTCCGCTCATTTCCCAAGGTGTTCTGCTGACGAGACGCTGCTCGATTC 18825
QY 611 TTCACGCCGCGCCGCCACACGAGCAGCGAATGATTTCCCTTCCATCTTCAGCTGATA 670
Db 18824 TTCACGCCGCGCCGCCACACGAGCAGCGAATGATTTCCCTTCCATCTTCAGCTGATA 18765
QY 671 CAATACACGAGCAGTAATTCATGCTCTTTTCGGGACGTAGCATCCCCACCTGAACGAT 730
Db 18764 CAATACACGAGCAGTAATTCATGCTCTTTTCGGGACGTAGCATCCCCACCTGAACGAT 18705
QY 731 AAGCGGAACATTTCTGCTGATGACGCCAGCGGTGGATATGACGGGTAAACGCTGCGAT 790
Db 18704 AAGCGGAACATTTCTGCTGATGACGCCAGCGGTGGATATGACGGGTAAACGCTGCGAT 18645
QY 791 GGCCTTCATATGCAATCGCGGCCAGTCGAAACCCGGTGGAAATACCGTTACCGGTGCT 850
Db 18644 GGCCTTCATATGCAATCGCGGCCAGTCGAAACCCGGTGGAAATACCGTTACCGGTGCT 18585
QY 851 GACACCTTCGCCCATCAGATGCCCATCATGGGTGAGATAGGCACAAATGAATACACA 910
Db 18584 GACACCTTCGCCCATCAGATGCCCATCATGGGTGAGATAGGCACAAATGAATACACA 18525
QY 911 CAGATAATTCAGGGAACGTTCTGCTTACGGGTGATGATAGGTTTTTGTCTGCAAT 970
Db 18524 CAGATAATTCAGGGAACGTTCTGCTTACGGGTGATGATAGGTTTTTGTCTGCAAT 18465
QY 971 AGTGAAGCGTGACAGCATATCAGCGCTGCTGCTGCTATATCTGTCATGCGGCAT 1030
Db 18464 ACTGAAGCGTGACAGCATATCAGCGCTGCTGCTGCTATATCTGTCATGCGGCAT 18405
QY 1031 ATGGCAGATGACAGATCAGGTTTAAATTCGCCGATAATCCGTCGAGTCTGAGGATGA 1090
Db 18404 ATGGCAGATGACAGATCAGGTTTAAATTCGCCGATAATCCGTCGAGTCTGAGGATGA 18345
QY 1091 AGGAAGTGAAGCTGTTCTGAAAGGAATAAAGTGACATATGCGCTTTTCTGCGC 1150
Db 18344 AGGAAGTGAAGCTGTTCTGAAAGGAATAAAGTGACATATGCGCTTTTCTGCGC 18285
QY 1151 TTCGGGAGCAATTTTACTTTTTTCTCTGCAG 1181
Db 18284 TTCGGGAGCAATTTTACTTTTTTCTCTGCAG 18254
```

```
RESULT 5
AX191727/c 92077 bp DNA linear PAT 15-AUG-2001
LOCUS Sequence 9 from Patent WO0149775.
DEFINITION AX191727
ACCESSION AX191727.1 GI:15209896
VERSION
KEYWORDS
SOURCE
ORGANISM
Escherichia coli.
Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
Escherichia coli.
REFERENCE
1 (bases 1 to 92077)
AUTHORS Iversen,P.L.
TITLE Antisense antibacterial cell division composition and method
JOURNAL Patent: WO 0149775-A 9 12-JUL-2001;
Avi Biopharma, Inc. (US)
FEATURES
Location/Qualifiers
1..92077
/organism="Escherichia coli"
/db_xref="taxon:562"
```

```
BASE COUNT 24593 a 20133 c 23728 g 23523 t
ORIGIN
Query Match 98.7%; Score 1166.2; DB 6; Length 92077;
Best Local Similarity 99.7%; Pred No. 0;
Matches 1188; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 11 TGGAAAAAGCCAAAATAAAAAATGCCCCATCCAGCGCTCCAGCTGAAAGTAGGCC 70
Db 19424 TGGAAAAAGCCAAAATAAAAAATGCCCCATCCAGCGCTCCAGCTGAAAGTAGGCC 19365
QY 71 TGTCTGTCCGGTATTAAATGCAATGACCGTCCCGTATTAAACAATGTGATAATTA 130
Db 19364 TGTCTGTCCGGTATTAAATGCAATGACCGTCCCGTATTAAACAATGTGATAATTA 19305
QY 131 CTCGGTTACCGGAAAAACCGCTGAACAAAATTCGGGCTGAAAAGAGATCCCGCTATCT 190
Db 19304 CTCGGTTACCGGAAAAACCGCTGAACAAAATTCGGGCTGAAAAGAGATCCCGCTATCT 19245
QY 191 GTTGCAATTTCCCTTACGCTGACTAGCAGACACAAATGATCTGTGCCGTCTCTGTAAT 250
Db 19244 GTTGCAATTTCCCTTACGCTGACTAGCAGACACAAATGATCTGTGCCGTCTCTGTAAT 19185
QY 251 ATCAAAACCGGTACTCAATATCTTCTGTGGGCTGGCTGCCATCATCGGAGCGTTCCGG 310
Db 19184 ATCAAAACCGGTACTCAATATCTTCTGTGGGCTGGCTGCCATCATCGGAGCGTTCCGG 19125
QY 311 TCGGGATAAAAAATTCGCGAGTCGCGGCTCCATGACAGACACATCCCCACGGGTAAACAG 370
Db 19124 TCGGGATAAAAAATTCGCGAGTCGCGGCTCCATGACAGACACATCCCCACGGGTAAACAG 19065
QY 371 CGTCCCTGTACATTTCTTGAATGACATCAGGGATCCCCCGCTCTCACTGGCGATAAC 430
Db 19064 CGTCCCTGTACATTTCTTGAATGACATCAGGGATCCCCCGCTCTCACTGGCGATAAC 19005
QY 431 GGCGACCGCGGAGACTGACGCTTCAGCAGTACCATACCAACGCTTCATTTCCGAGG 490
Db 19004 GGCGACCGCGGAGACTGACGCTTCAGCAGTACCATACCAACGCTTCATTTCCGAGG 18945
QY 491 CATGACACACACTGGCAATCGGTAGACCGTAAACGCTGGGAAAGGACCTGCGCAT 550
Db 18944 CATGACACACACTGGCAATCGGTAGACCGTAAACGCTGGGAAAGGACCTGCGCAT 18885
QY 551 TAACACATCTCCGCTCATTTCCCAAGGTGTTCTGCTGACGAGACGCTGCTCGATTC 610
Db 18884 TAACACATCTCCGCTCATTTCCCAAGGTGTTCTGCTGACGAGACGCTGCTCGATTC 18825
QY 611 TTCACGCCGCGCCGCCACACGAGCAGCGAATGATTTCCCTTCCATCTTCAGCTGATA 670
Db 18824 TTCACGCCGCGCCGCCACACGAGCAGCGAATGATTTCCCTTCCATCTTCAGCTGATA 18765
QY 671 CAATACACGAGCAGTAATTCATGCTCTTTTCGGGACGTAGCATCCCCACCTGAACGAT 730
Db 18764 CAATACACGAGCAGTAATTCATGCTCTTTTCGGGACGTAGCATCCCCACCTGAACGAT 18705
QY 731 AAGCGGAACATTTCTGCTGATGACGCCAGCGGTGGATATGACGGGTAAACGCTGCGAT 790
Db 18704 AAGCGGAACATTTCTGCTGATGACGCCAGCGGTGGATATGACGGGTAAACGCTGCGAT 18645
QY 791 GGCTTCATATGCAATCGCGGCCAGTCGAAACCCGGTGGAAATACCGTTACCGGTGCT 850
Db 18644 GGCTTCATATGCAATCGCGGCCAGTCGAAACCCGGTGGAAATACCGTTACCGGTGCT 18585
QY 851 GACACCTTCGCCCATCAGATGCCCATCATGGGTGAGATAGGCACAAATGAATACACA 910
Db 18584 GACACCTTCGCCCATCAGATGCCCATCATGGGTGAGATAGGCACAAATGAATACACA 18525
QY 911 CAGATAATTCAGGGAACGTTCTGCTTACGGGTGATGATAGGTTTTTGTCTGCAAT 970
Db 18524 CAGATAATTCAGGGAACGTTCTGCTTACGGGTGATGATAGGTTTTTGTCTGCAAT 18465
QY 971 AGTGAAGCGTGACAGCATATCAGCGCTGCTGCTGCTATATCTGTCATGCGGCAT 1030
```

```

Db      18464  ACTGAACGGTGACAGCATATCAGACGGCTCAGTCTGCTGATATTAATCTACTGTCATGGCCACT 18405
Qy      1031  ATGCAGATGACCATGACGATGAGTTAAATCCCGATAATCCGTCGAAGTCTGAGGATGGA 1090
Db      18404  ATGCAGATGACCATGACGATGAGTTAAATCCCGATAATCCGTCGAAGTCTGAGGATGGA 18345
Qy      1091  AGGAAGTGAGGCTGTTCTCTGAAAGGAATAAAGTGACATCATGCCCTCTTTTCTGGC 1150
Db      18344  AGGAAGTGAGGCTGTTCTCTGAAAGGAATAAAGTGACATCATGCCCTCTTTTCTGGC 18285
Qy      1151  TTCGGAGCAATTTTACTTTTTTCTCTGCAG 1181
Db      18284  TTCGGAGCAATTTTACTTTTTTCTCTGCAG 18254

RESULT 6
LOCUS      AB011549          92721 bp    DNA    circular BCT 27-APR-1999
DEFINITION Escherichia coli plasmid pO157 DNA, complete sequence.
ACCESSION  AB011549
VERSION    AB011549.2 GI:4589740
KEYWORDS   ToxR-regulated lipoprotein; tagA.
SOURCE     Escherichia coli (strain:O157:H7, sub_strain:RIMD 0509952)
           plasmid:pO157 DNA.
ORGANISM   Escherichia coli
           Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
           Escherichia.
REFERENCE  1 (sites)
AUTHORS   Makino,K., Ishii,K., Yasunaga,T., Hattori,M., Yokoyama,K.,
           Yutsudo,H.C., Kubota,Y., Yamaichi,Y., Iida,T., Yamamoto,K.,
           Honda,T., Han,C.G., Ohtsubo,E., Kasamatsu,M., Hayashi,T., Kuhara,S.
           and Shingagawa,H.
TITLE      Complete nucleotide sequences of 93-kb and 3.3-kb plasmids of an
           enterohemorrhagic Escherichia coli O157:H7 derived from Sakai
           outbreak
JOURNAL   DNA Res. 5 (1), 1-9 (1998)
MEDLINE   98290340
REFERENCE  2 (bases 1 to 92721)
AUTHORS   Makino,K.
TITLE      Direct Submission
JOURNAL   Submitted (24-FEB-1998) Kozo Makino, Research Institute for
           Microbial Diseases, Osaka University, Molecular Microbiology;
           Yamadaoka, 3-1, Suita, Osaka 562, Japan
           (E-mail:makino@bks01.biken.osaka-u.ac.jp, Tel:81-6-879-8318,
           Fax:81-6-879-8320)
COMMENT   On Apr 20, 1999 this sequence version replaced gi:3336997.
FEATURES   Location/Qualifiers
           1..92721
             /organism="Escherichia coli"
             /strain="O157:H7"
             /sub_strain="RIMD 0509952"
             /db_xref="taxon:562"
             /plasmid="pO157"
             /note="RIMD 0509952 is a strain of enterohemorrhagic E.
             coli, EHEC O157:H7"
           join(92527..92721,1..2502)
             /gene="tagA"
           join(92527..92721,1..2502)
             /gene="tagA"
           /codon_start=1
           /transl_table=11
           /product="ToxR-regulated lipoprotein"
           /protein_id="BAA31757.3"
           /db_xref="GI:4666293"
           /translation="MNTKMERWTPMKLYLSCTILAPLAGVSTADNNSAIVE
           NTSQINDLQSLAAEVKFAQSQILPAHPKGSQSHLSRLSLILVRPVKADKTP
           VQVARDNKHSLTLTLPSPSLPTIYHLDVPEGGIDFTPHNGTKLIINTVAEYN
           KLSADSSSHLTHNALVEIHTANGRWYRDYLPQGPDLGKMYRFVSSAGYSTY
           FYGDKVTLTGVLFFKYGWQFRSGELENRTIYAQIWSAELPAHWIYVPLNLV
           IKQNLGRINDIKIGAPGELLHTIDIGMLTTPDRDFDFAKDCAHREYFQIPYSR
           MIVNYPALHLKVMYLPGLTMDPGNGWHSQGMORIGKELYSHGIDNANYGLN
           STAGLGENSPHYVAQLAAHNSRGNLANGLVHGGGSGGIVILDSITNEKSFHEVGH
           NYGLGHVYDGFKGSVHRSABNNNSNTGWGDGKKRFPINFPYPSQTNKSLNNOCQEPF

```

```

DGHKGFDMAGGSPFSANRFTWTTPNSSAIIOREFFENKAVFDSRSTGFSKKNADT
QEMPEYHTIDRAEQITASVNELSESKMAELMAEYAVYVKVHMWNGNFTNRYITPASA
DNMGSLITINHEAGYSYLVINGDESKYVSGYKKSFSVSDGQKRWEDYVTRREARKPE
QFQVPTLVLYGVDPEGLSSYIYPMATGYGFTYSDSONLSDNCOLOVDTTKEGQL
REFLANHRANTVMNKFHINVTESQPTQATLVCKNKLIDTKSLTPAPEGLTYTVNGO
ALPAKENGCCIVSNKSRKRYCLPVQORSYSLPDWIVQGEYVYVDSGAKAKVLLSDWDN
LSYNRIGBFGVGNVNPADMKKVKAMNGQYLDPSKPSMRKVVYK"
2589..3464
/gene="etpC"
2589..3464
/gene="etpC"
/codon_start=1
/transl_table=11
/product="Type II secretion pathway related protein"
/protein_id="BAA31758.1"
/db_xref="GI:3336999"
/translation="MLFSLSPGRDGLFKDVLKMLTPNRLCLVILLIAGYQLVSVI
HFHWTQAASVPLSKRSVAPETAVTGDQTEERFVFLFGRASPLSSEGRAGETMPSLS
DQLSGEDLVIRGILYSVAEHSVAIFAHNRRQFSLSGEKVPSYDATISAIFSDHIY
INYGKTVSLPLRYDNTKKNAYDNNNLTVGDVITQDNFRVESVFDIMSFSAVTYNT
LSGYRLIPGKHSSLYNAGLHNDLAVSVNGSELDRTRAQOIMKQLPKELKEITVE
RDGQLYDAFIAVGEN"
3675..5432
/gene="etpD"
3675..5432
/gene="etpD"
/codon_start=1
/transl_table=11
/product="Type II secretion pathway related protein"
/protein_id="BAA31759.1"
/db_xref="GI:3337000"
/translation="MLNEEQYYQFFLSVLDVYGFVAVDMHNGILKVVRSKDAKTSAPV
VADSVPGTGDEVTWVVPVSVNAARDLAPLRLDNLNAGAGSVHYEYPSNVLLMTGR
AASMKRLMEIVERDKVGNRSVATVPLTASATDVARLVLTETKTDTATLPAWMTAK
LYADERTNSLVSGEPISQORTIISIIKOLDQEDVQGNTKVYLYKAKADLDELVTG
ISSIENDSKSPSTEALARKGVYIKSHQTNALLITGAPDVIKRDLENVLSQIDIRPQ
VLVEALIAIQDADGILNIGIQWNRHAGVAQTGLPTTMTVQTRQNEILDSQNSA
LSMFGNIAAGFYQGNWMLLTALSTSSKNILATPSIVTLDMNEATFNVQGEVPLVG
SQTTSQDNIFNTVERKTVGILKRVKQINEGSVLLEIEQEVSGVADTAVATTDLGA
TENTRTVNAMLVNGEIVVVGGLDKSIRGSESKVPLIGDIPVGLHLPKASEQTAK"
RNLMLFRTIIRERDGRFHAESAERYQSEFNQGVQSRGKETTALTTLNEQLRLSPDQD
DPAFRKVAIAAFYAEQEM"
5432..6937
/gene="etpE"
5432..6937
/gene="etpE"
/codon_start=1
/transl_table=11
/product="Type II secretion pathway related protein"
/protein_id="BAA31760.1"
/db_xref="GI:3337001"
/translation="MSRVQNVSESRPLLPFSFRTQRIILLIREQGNRVFCMEDTPA
SALLFVRVAEGPLNVTTVSAEAFKQLVSSYORDSDDEAROMMAEIGNEMDPTVAGE
LPDREDLDANDDAPILRLINAMLTAEAKESADJHETRYERHLQVRFIDGVLREIL
RUHRLNLSLISRIKVMKLDIAERKVPQDGRWVLRIGRAVDVRYVSTPSNHGRIY
LRLLDKNSVSLDAAALGQQNQOHIDALIRPHGIIIVTGTSGKSTLYAALSLL
NPRDRNIMTVEDEVYELDGIISQTVNPKVDMTFAFSLRAILRQDVPVLYGVEIRDE
TAQIAVOASLTGHLVLSLTHNSAAGSLRQMDGIPPELSTSLSLAVLAQLVRLTC
PRCROPQVSTELAMDMDIPPETTIWOPAGQCHCSFTGVGHGTGTHEILLIDIRTA
IYQGEIGELGITRACGRYLTLRDGRQKVLACETSWEEVVRVVTESRLQEE"
6939..8162
/gene="etpF"
6939..8162
/gene="etpF"
/codon_start=1
/transl_table=11
/product="Type II secretion pathway related protein"
/protein_id="BAA31761.1"
/db_xref="GI:3337002"
/translation="NALFYQASDIHGRKRSIGLEADSARHARQLLREQALIPVRIDE
KOVHHKHSRSTILRFRPGSSAEALALTROLATLVAASLPLEEALDALLROSEKPRQ
RNLIAAVRTVLEGHSLAAMGMFQTFERLVCAMVAGSETSRGLDVLRLADYEQ
RLIMRNLRLQALLPCVLTLVAVGTAILLTAVPVKVEQFIHMKQTLPLSLVLMGA

```


Db 87919 ATGGCAGATGACCATGAGTTCATTAATCCCGAATAATCCGTCGAAGTCTGAGGATGGA 87860
QY 1091 AGGAAGTGAAGGCTGTTCTTGAAGGAATAAAGTACATCATGCCCCCTTTTCTGGC 1150
Db 87859 AGGAAGTGAAGGCTGTTCTTGAAGGAATAAAGTACATCATGCCCCCTTTTCTGGC 87800
QY 1151 TTCGGGAGCAATTTTACTTTTCTGTCGAG 1181
Db 87799 TTCGGGAGCAATTTTACTTTTCTGTCGAG 87769
RESULT 7
AX191725/c
LOCUS AX191725 92721 bp DNA linear PAT 15-AUG-2001
DEFINITION Sequence 7 from Patent WO0149775.
ACCESSION AX191725
VERSION AX191725.1 GI:15209894
KEYWORDS
SOURCE Escherichia coli.
ORGANISM Escherichia coli.
Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
Escherichia.
REFERENCE 1 (bases 1 to 92721)
AUTHORS Iversen, P. L.
TITLE Antisense antibacterial cell division composition and method
JOURNAL Patent: WO 0149775-A 7 12-JUL-2001;
Avi Biopharma, Inc. (US)
FEATURES
source Location/Qualifiers
1..92721
/organism="Escherichia coli"
/db_xref="taxon:562"
BASE COUNT 24870 a 20267 c 23868 g 23716 t
ORIGIN
Query Match 98.7%; Score 1166.2; DB 6; Length 92721;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1168; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 11 TGGAAAAAGCCAAATATAAAAAATGCCCCATCCCGCGCTCCAGCTGAAAGTAGGCC 70
Db 88939 TGGAAAAAGCCAAATATAAAAAATGCCCCATCCCGCGCTCCAGCTGAAAGTAGGCC 88880
QY 71 TGTTCCTGCGGTATTTAAATGCAATGACCGTCCCGCTATTTAAACAAATGTGATAATTA 130
Db 88879 TGTTCCTGCGGTATTTAAATGCAATGACCGTCCCGCTATTTAAACAAATGTGATAATTA 88820
QY 131 CTCGGTTACCGGAAACCGCTGAACAAAATTCGGCTGAAAGAGGATCCCGCTTATCT 190
Db 88819 CTCGGTTACCGGAAACCGCTGAACAAAATTCGGCTGAAAGAGGATCCCGCTTATCT 88760
QY 191 GTTGCAATTTCCCTTAGCCTGACTAGCCAGACACAAATGATGTCGCGGTTCTGTTAAT 250
Db 88759 GTTGCAATTTCCCTTAGCCTGACTAGCCAGACACAAATGATGTCGCGGTTCTGTTAAT 88700
QY 251 ATCAACCGGTACTCAATATCTTCTGCGCTGGCTGCCATCATCCGGAAGCGTCCGG 310
Db 88699 ATCAACCGGTACTCAATATCTTCTGCGCTGGCTGCCATCATCCGGAAGCGTCCGG 88640
QY 311 TCGGATATAAATCGCGAGTGCCTGATCATGACACACATCCCGCCAGCGGTAAACAG 370
Db 88639 TCGGATATAAATCGCGAGTGCCTGATCATGACACACATCCCGCCAGCGGTAAACAG 88580
QY 371 CTTCCCTGTCAATTTCTTGAATGACATCAGGATCCCGCGCTCTCAGTGGCGATTAAC 430
Db 88579 CTTCCCTGTCAATTTCTTGAATGACATCAGGATCCCGCGCTCTCAGTGGCGATTAAC 88520
QY 431 GGGCAGCGGAGACTGAGCTTACCGAGTACCATACCAACCGTTCATTTTCGGAAGG 490
Db 88519 GGGCAGCGGAGACTGAGCTTACCGAGTACCATACCAACCGTTCATTTTCGGAAGG 88460
QY 491 CATGACACACATGGCAATCCGGTAGACCGGTAAACGCTGGGAAAGGCGACCTGGCCAT 550
Db 88459 CATGACACACATGGCAATCCGGTAGACCGGTAAACGCTGGGAAAGGCGACCTGGCCAT 88400

551 TAACACATCTCCGCTCAATCCAGGTGTTCTCTCTGCTGACGACGACGTGCTTCGTATTC 610
88399 TAACACATCTCCGCTCAATCCAGGTGTTCTCTCTGCTGACGACGACGTGCTTCGTATTC 88340
QY 611 TTCAGCCCGCGGCCACACAGGAGCAAGGAATGATTTCCCTCCATCTTCAGCTGATA 670
Db 88339 TTCAGCCCGCGGCCACACAGGAGCAAGGAATGATTTCCCTCCATCTTCAGCTGATA 88280
QY 671 CAATACACGACGACATAAATTCATGCTCTTTTCGGGACGTAGCATCCCACTGAACGAT 730
Db 88279 CAATACACGACGACATAAATTCATGCTCTTTTCGGGACGTAGCATCCCACTGAACGAT 88220
QY 731 AAGCGGAACATTTCTGCTGATGTCAGCCAGCGGTGGATATGACGGGTGAAGCGTCGAT 790
Db 88219 AAGCGGAACATTTCTGCTGATGTCAGCCAGCGGTGGATATGACGGGTGAAGCGTCGAT 88160
QY 791 GCGTTCAATTCATGCGGCCAGTCGAAACCGGTGGAATAACCGTTACCGGTGCTCT 850
Db 88159 GCGTTCAATTCATGCGGCCAGTCGAAACCGGTGGAATAACCGTTACCGGTGCTCT 88100
QY 851 GACACCTTCGCGCATCAGATGCGGCATCATGCTGAGATAGGACACAAATGAAATCACA 910
Db 88099 GACACCTTCGCGCATCAGATGCGGCATCATGCTGAGTAGGACACAAATGAAATCACA 88040
QY 911 CAGATAAATTCAGGAAACGTTCTGCTCTTACGGGTGATGAGTTTCTGCTGACAAAT 970
Db 88039 CAGATAAATTCAGGAAACGTTCTGCTCTTACGGGTGATGAGTTTCTGCTGACAAAT 87980
QY 971 AGTGAAGCGGTGACACATATCAGACGCTCAGTCTCTGCTATATATCTCATGCGCACT 1030
Db 87979 ACTGAAGCGGTGACACATATCAGACGCTCAGTCTCTGCTATATATCTCATGCGCACT 87920
QY 1031 ATGCGAGATGACGACATCAGTTTAAATTCGCCGATAATCCCGTAAGTCTGAGGATGA 1090
Db 87919 ATGCGAGATGACGACATCAGTTTAAATTCGCCGATAATCCCGTAAGTCTGAGGATGA 87860
QY 1091 AGGAAGTGAAGGCTGTTCTTGAAGGAATAAAGTACATCATGCCCCCTTTTCTGGC 1150
Db 87859 AGGAAGTGAAGGCTGTTCTTGAAGGAATAAAGTACATCATGCCCCCTTTTCTGGC 87800
QY 1151 TTCGGGAGCAATTTTACTTTTCTGTCGAG 1181
Db 87799 TTCGGGAGCAATTTTACTTTTCTGTCGAG 87769
RESULT 8
AF043470
LOCUS AF043470 5612 bp DNA linear BCT 25-JUN-1998
DEFINITION Escherichia coli plasmid pO157 ecf4 gene, partial cds; and ecf3, ecf2, and ecf1 genes, complete cds.
ACCESSION AF043470
VERSION AF043470.1 GI:3253288
KEYWORDS
SOURCE Escherichia coli.
ORGANISM Escherichia coli.
Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
Escherichia.
REFERENCE 1 (bases 1 to 5612)
AUTHORS Boerlin, P., Chen, S., Collbourne, J.K., Johnson, R., De Grandis, S. and Gyles, C.
TITLE Evolution of enterohemorrhagic Escherichia coli hemolysin plasmids and the locus for enterocyte effacement in shiga toxin-producing E. coli
JOURNAL Infect. Immun. 66 (6), 2553-2561 (1998)
MEDLINE 98261495
PUBMED 9596716
REFERENCE 2 (bases 1 to 5612)
AUTHORS Boerlin, P. and Gyles, C.
TITLE Direct Submission
JOURNAL Submitted (17-JUN-1998) Pathobiology, ON Veterinary College, University of Guelph, Guelph, ON N1G 2W1, Canada
FEATURES Location/Qualifiers

source	1. .5612 /organism="Escherichia coli" /strain="4304-Pr14" /serotype="O157:H7" /db_xref="taxon:562" /plasmid="p0157" /note="produces Shiga toxin" complement(<1. .705) /gene="ecf4" complement(<1. .705) /gene="ecf4" /codon_start=1 /transl_table=11 /product="unknown" /protein_id="AAC24348.1" /db_xref="GI:3253289" /translation="MVPPSAVLCYHNEISROIQIPVNMKNIRTEFIPRNFNLTPFPYWM TWGICGVFAMVPPALRDLGLKMLVGRKLSARORALINLSLCPPEYSDREKE NIVDAMFATASMAVLMALALSGDPKISHRIWNGLEIVEKMAQNNKVIPLVPHAW GYDIPAMLAASGRKMAFMHQRNPVVDVYVNSVRRREGGKLHARDGIASFVRSVR QGWYGLPDDQHGPEF" complement(751. .2499) /gene="ecf3" complement(751. .2499) /gene="ecf3" /codon_start=1 /transl_table=11 /product="unknown" /protein_id="AAC24349.1" /db_xref="GI:3253290" /translation="MHLNLTQNPFTSWALGWAIFYFGFFFTLLQVIFSSGYSGTN GKDSLFLSCLMIPFLYDRIKIIAAVVGFLLMGTSLAALCYIFLYGHEFSQVLF VMEFNARAGSFYSQVLSKLLLSLAVTAVSVFLWTLRPVYIPLWRRIVSFLLL YALLPWLKSLIROEPLNDLTKLASMEPAAPQWFSVSYQYHQOALNLTFLNE NSALPLGLNDESGRPRTLVLIGESTORERMSLYGLRETTPDLALRKTDPGLT VFNVVASPYTEALQOALTANENKPDLYLTOPSLMMKQAGYKTFWITNQOIT ARTMLTVSRQTRDQYVYMNQRTQSAREYDYNLPFREYLDNPAPKLLIVHLLGT HKIKRYPEGGRFGTIGTIPITGLNAKELEYINDYDANLFDHNVASLIDKFRAT APGFLLYSDGHEEYDTPPYKTQGRNEDNTPRMYTPVFLWTSEKHAHPRDFS QYDRIYSLAELIHTWSDLAGLYDGYDTPTRSLVNPQFRETTRWIGNPYKKNLITD LFRVSHSIFLSSVPP" complement(2589. .3695) /gene="ecf2" complement(2589. .3695) /gene="ecf2" /codon_start=1 /transl_table=11 /product="unknown" /protein_id="AAC24350.1" /db_xref="GI:3253291" /translation="MMKILFTSSSDIGGOELQALQMTALQKQHSVLLACREKSKI APARRGHDVTFIPRNSHLPSILFRRIIGEFKPDVIVCHSGHDSNIAGLSLRIC CHRSIVROKTYITRKTTFSLNYLQDFIVPSSAMHLMAGVTRPTVTVIPPGDW PALHNGAMRPLHIIHAAASADNVLIVQVGMRLPEKGEHEMLRYLYOLKMGKSR WLVVAGREYERASLQQOEHGLMSGDLVLMAGALFPALPVRIASVYVWPSNEAFGM VLAASVSGVPVIASTGGIPDVIOKNVTGTLPLPGVDVSAWNTGALRDLRSRPERFRM AASAREDIYRFDINRTAQIIVLSAQAQKCNR" complement(3695. .4516) /gene="ecf1" complement(3695. .4516) /gene="ecf1" /codon_start=1 /transl_table=11 /product="unknown" /protein_id="AAC24351.1" /db_xref="GI:3253292" /translation="MLNARHLPVLMYHHVSRCPGLVTLSPETCEQMKWLADNNRTV TSAELFYQGGFLPKRSVMLTFDDGLDNLWLRAPVLYEINLHAIFLITGLIGKN VRSRQEQSHRCELIADNRSDVLRWSEVRMRDLSGLVEFHVHTSHKRWDRLS VSRACQRLMKEDILVQKCLTEKLGFCSSHLCPWEPGYNRYDINLAGLKGFSLYTT ERMNCPENGLSRIGISTKERHSGWLKRLRYTYTTPLFSSVLAHLKGPRLPDN" complement(1574 a 1399 c 1207 g 1432 t
gene	11 TGGAAAAAGCCAAAAATAAAATGCCCCATCCAGCGCGCTCCAGCTGAAAGTAGGCC 70 2415 TGGAAAAAGCCAAAAATAAAATGCCCCATCCAGCGCGCTCCAGCTGAAAGTAGGCC 2474
CDS	71 TGTTCGTCCGGTATTAAATGCATTGACCGTCCCGGTATTTAAACAATGTGATAAATTA 130 2475 TGTTCGTCCGGTATTAAATGCATTGACCGTCCCGGTATTTAAACAATGTGATAAATTA 2534
gene	131 CTCCTGTACCGGAAACCGCTGAACAAAATTGGGGCTGAAAAGAGGATCGCGGTATCT 190 2535 CTCCTGTACCGGAAACCGCTGAACAAAATTGGGGCTGAAAAGAGGATCGCGGTATCT 2594
gene	191 GTTGCATTTCCTTACCTGACTGACGAGACACAATGATCTGTGCGTCTCTGTTAAT 250 2595 GTTGCATTTCCTTACCTGACTGACGAGACACAATGATCTGTGCGTCTCTGTTAAT 2654
gene	251 ATCAAAACGGTACTCAATATCTTCTGGCGTGGCTGCCATCATCCGGAAGCGTTCGG 310 2655 ATCAAAACGGTACTCAATATCTTCTGGCGTGGCTGCCATCATCCGGAAGCGTTCGG 2714
gene	311 TCGGGATAAAAATCGCGGAGTGGCGGTCCATGACAGACACATCCCCACAGGGGTACAG 370 2715 TCGGGATAAAAATCGCGGAGTGGCGGTCCATGACAGACACATCCCCACAGGGGTACAG 2774
gene	371 CGTCCCTGTACATCTTCTGAATGACATCAGGATCCCGCGCTCTCACTGGCGATAAC 430 2775 CGTCCCTGTACATCTTCTGAATGACATCAGGATCCCGCGCTCTCACTGGCGATAAC 2834
gene	431 GGCACGCGGGAGACTGAGCTTCAGCCAGTACCATACCAACAGCTTCATTTCCGAAGG 490 2835 GGCACGCGGGAGACTGAGCTTCAGCCAGTACCATACCAACAGCTTCATTTCCGAAGG 2894
gene	491 CATGACACACACTGGCAATCGGTAGACCGGTAAAGTGGGAAAGGACCGTCCGAT 550 2895 CATGACACACACTGGCAATCGGTAGACCGGTAAAGTGGGAAAGGACCGTCCGAT 2954
gene	551 TAACACATCTCCGCTCATTTCCAGGTGTTCTGCTCTGACGACGAGTCTTCGTATTC 610 2955 TAACACATCTCCGCTCATTTCCAGGTGTTCTGCTCTGACGACGAGTCTTCGTATTC 3014
gene	611 TTCACGCGCGCGCCACACAGCCAGCAATATGATTTCCCTTCCATCTTCAGCTGATA 670 3015 TTCACGCGCGCGCCACACAGCCAGCAATATGATTTCCCTTCCATCTTCAGCTGATA 3074
gene	671 CAATACACGAGCATAAATTCATGTCCTTTTCGGGACGTAGCATCCCACTGAAGCAT 730 3075 CAATACACGAGCATAAATTCATGTCCTTTTCGGGACGTAGCATCCCACTGAAGCAT 3134
gene	731 AAGCGAAACATTTCTGCTGATGCAGCCAGCGTGTATGATGAGGGGTAAACGGTCCGAT 790 3135 AAGCGAAACATTTCTGCTGATGCAGCCAGCGTGTATGATGAGGGGTAAACGGTCCGAT 3194
gene	791 GGCTTCATTTATGCAATCGCGGCGCAGTCGAAACCCGGTGGAAATACCGTTACCGGTGCT 850 3195 GGCTTCATTTATGCAATCGCGGCGCAGTCGAAACCCGGTGGAAATACCGTTACCGGTGCT 3254
gene	851 GACACCTTCGGCATCAGATGCGCCATCATGGGTGAGATAGGCACAAATGAATCACA 910 3255 GACACCTTCGGCATCAGATGCGCCATCATGGGTGAGATAGGCACAAATGAATCACA 3314
gene	911 CAGATATTCAGGAAACCGTCTGCTTACGGGTGATGTAGGTCTTTTCTGTGCAAT 970 3315 CAGATATTCAGGAAACCGTCTGCTTACGGGTGATGTAGGTCTTTTCTGTGCAAT 3374
gene	971 AGTGAAGCGGTGACAGCATATCAGACGGCTCAGTCTCTGCTATATTACTGTGATGGCACT 1030 3375 ACTGAAGCGGTGACAGCATATCAGACGGCTCAGTCTCTGCTATATTACTGTGATGGCACT 3434


```

/codon_start=1
/product="ospf, secreted by the Mxi-Spa secretion
machinery, function unknown"
/protein_id="CAC05773.1"
/db_xref="GI:12329042"
/translation="MPIKPKLNLDSLNVKSEIPOMLSANERLKNFNILYNOIR
QPAYPKVASNPVTSDICQFVSMYQGFQIVNHSGDVFHACRNPSKSGDFVGDK
FHSIAREQVLAQILSGLLSESDPIDKWKITDMNRVSOQSRVIGIAQFTLYKSD
QECQSYALLHRIQFIMCLESNLLRSKIAPEYSPASVDRPEDMKYVSYRNLRSR
DSEROEQLREPFYRLMIE"
/complement(13393..13971)
/notes="orf13, length= 192 aa, unknown"
/codon_start=1
/transl_table=11
/protein_id="CAC05774.1"
/db_xref="GI:12329043"
/translation="MKVPSKSLGYIFHDYINKKHTIDFNDVVRKAVLSGKINELNAC
HKVAIFLAKDNEITKKDKAKIIDTLTENYSIEFQOLMNISETLNSLYITPESGF
VSFVNEGKICHYAYKSDNSMTYHANGSSIDKXITDMCGLICMRHIESIGLIIFM
LDKVLISAETEFMNERGWRAPFCASAKNLKYCV"
14014..14241
/notes="IS2.02, 90% id over 228 nt with IS2, from 4 to 230"
/complement(14367..14629)
/notes="IS21.01, 97% id over 263 nt with IS21, from 1825 to
2098"
complement(14630..15558)
/notes="IS91.01, 97% id over 929 nt with IS91, from 191 to
0"
15559..15924
/notes="ISshf5.02a, id to ISshf5 from 1 to 366"
complement(15925..17350)
/notes="IS4.01, 99% id over 1426 nt with IS4, from 1 to 0"
17351..19039
/notes="ISshf5.02b, id over 1689 nt with ISshf5, from 335
to 0"
19326..19523
/notes="IS3.01, 88% id over 198 nt with IS3, from 1062 to
0"
20459..20601
/notes="IS630.02, 84% id over 143 nt with IS630, from 999
to 1141"
20964..21641
/genes="ospd1"
20964..21641
/genes="ospd1"
/notes="ORF21, length = 225 aa, 34.3% id in 140 aa overlap
with BCSHET2B.1 (S.flexneri SenA - 565 aa)"
/codon_start=1
/transl_table=11
/product="ospd1, secreted by the Mxi-Spa secretion
machinery, function unknown"
/protein_id="CAC05775.1"
/db_xref="GI:12329044"
/translation="MSINNVGLHPANNKWHLLIGSTANENKMKNNIINVTNFAIS
HAINRKSGGCGVGFVKLAKIONISIPKNNKYNRNHLFSLIWHGNADARKYSE
SLUAEIPEERLEVLNARNAGESALFIALQEGHSAAYOAGDFIKTFDLSPKETIK
LLDVRDNEGLPLAAGRGNIEMMAYINICHHSIGIKLIEIADRLNNNEQDMFNIS
DKIQELF"
21977..22105
/notes="IS3.02, 89% id over 129 nt with IS3, from 1130 to
0"
22153..22467
/notes="ORF22, length = 104 aa, unknown"
/codon_start=1
/transl_table=11
/protein_id="CAC05776.1"
/db_xref="GI:12329045"
/translation="MLQRGRKVGFAQLPVDFVAIEPDSVQGVGKRLNTRCFIIRI
27.8%; Score 328.6; DB 1; Length 213494;
Best Local Similarity 59.3%; Pred. No. 1.4e-94;
Matches 578; Conservative 0; Mismatches 394; Indels 3; Gaps 1;

```

210 TGACTAGCCAGACACACAATGATCTGTGCCGTTCTGTTAATATCAAAACCGGTACTCAATA 269
 187295 TGCTTCGCTAATGTGAGTATTTTAAACGCAAGTTTATATATCAAAACCGCTCTCTATA 187236
 270 TCCTTCCTGCGCTGGCTGCATCCCGAAGCGTTCGCGTGGGATAAAATACCGC 329
 187235 TCCTGTTTTCCTGGGAGCCATCTGATAAAACGCCCGGAGTCAATTAATAAATATTC 187176
 330 AGTCCGCGGTCCATGACAGACATCCCCACGCGGTAAACAGCGTCCCTGTGCATATCTTC 389
 187175 AGGCGCACATCCATCGCTGCTTATTACCTGCTGGTAACAATGTCCTGGTCTGGTTATTC 187116
 390 TGAATGACATCAGGATCCCGCCGCTCACTGCGGATACAGGCGCGGAGACTGAC 449
 187115 TGAATAACATCAGGATTCACCAATCTGACTGCCAGTAGACGACAGAAAATGCCGAT 187056
 450 GCTTCAGCCAGTACCATAACAAACGCTTCAATTTTCGAAGCATGACACACACTGGCA 509
 187055 GCTTCGCCAGTACCATAACAAACGATTCGTTTTCTGAAGGAGAGAACCCAGACTGGCA 186996
 510 ATCCGCTAGACCGGTACGCTGGGAAAGGCGACCTGCCATTAAACACATCTCCGCTCAT 569
 186995 ACCGATATACGGGGCGGCGAGAAAACATTTGCAGCAATAAAACATCATCATGATG 186936
 570 CCAGGTGTTCTGCTGCTGACGAGAGAGTGTCTGATTTCTTCAACCCCGGCGCCACC 629
 186935 CCATATCTGCTCAATCTGATCTGTAATGCTCCCGCAGTTTCAAGGCAACGAGACCCACA 186876
 630 ACCAGCAGCGAAATGATTTCCCTTCATCTTCAGCTGATACATACACGACGACATAAT 689
 186875 ATCAACCAACAGAACTGCTGCTCATTTTAAATGGAAACAGTAAATTTCAACATAAT 186816
 690 TCATGCTCTTTTCGGGACGTAGCATCCCAACCTGAAAGGAGGAGCAATTTCTCTGCT 749
 186815 TCCTGCTCTTTTCGGGCGCAACATTTCTACCTGAGCAATAAGCGGCGATCCCTTCGG 186756
 750 GATGACGCCAGCGGTGGATATGACGAGGGTAACGGTCCGATGGTTCATTTATGCAATGCG 809
 186755 GACGCCAGCGCAAGAAAGAACATTTGGAGGCAAACTGTTTCGGGAATCAAGTATAATTC 186696
 810 GGCCAGTCGAAACCGGTGGAATAACCGTTACCGGTGCTGACACCTTCGCCCATCAGA 869
 186695 TGAATCAAGCCTGGAGGCAACACAGTAAACCGGTTTCGACATCTCTTCTGCTCCAGA 186636
 870 TGGCCCATCATGGGTGAGATAGGCACAAACAAATGAATACACAGATAATTCAGGAGAAAC 929
 186635 TGTGCTCTTACTTGTTCGGGGAACAATCACCTCATCGCAAAATGATTTATTGAAAAA 186576
 930 GTTCTGCTTACGGGTGATGATAGGTTTTTTTGTCTGACAATAGTGAAGCGGTG---ACAG 986
 186575 ACTTTTGTGTTTTCGTGCAAAATATGCTTTTGTGCTGATAATTTGAAATGGATGTTTCTTA 186516
 987 CATATCAGACGGCTCAGTCCCTGCTATATTACTGTCAGGCGCACTATGGCAGATGACACAGA 1046
 186515 GTGAATAACCGCTACTAAACCAACAATATGCTATCATGCCCACTGTGACAAAGATGCA 186456
 1047 TCAGGTTTAAATTTCCCGATAATCCCGTAGTCTGAGGATGGAAGGAGGTGAAGCTG 1106
 186455 TTGGCTGAAACCATCAACTATTCGAGTAATCTCCATCCAGTAGGATGTGAGACTG 186396
 1107 TCCCTGAAAGGAATAAAGTGACATCATGCCCTTTTCTGCTTTTCTGGCTCCCGAGCAATTTTA 1166
 186395 TTTTCGAAATAACGGAATGTGATCAATCCCAATTTACTGGCTTCAAAAGCAATTTTG 186336
 1167 CTTTCTCTCTGCGAG 1181
 186335 CTGTTTTCTCTGCGAG 186321

RESULT 10
 AF386526/c AF386526
 LOCUS 221618 bp DNA circular BCT 01-FEB-2002


```

/codon_start-1
/transl_table-11
/product="hypothetical protein"
/protein_id="AAL72473.1"
/db_xref="GI:18462701"
/translation="MVHATPELAKHASSPCGWDFVEPKNTAVRSPESNRIAKSFVKTIK
CDYISMPADGLTAANKLAFAFEHYNHWPHSALDYSREYLRQRANDRCLEI"
485. .949
/gene="CP0002"
485. .949
/gene="CP0002"
/notes="similar to a 333 aa protein from Sphingomonas sp.
gb: RAD45420.1"
/codon_start-1
/transl_table-11
/product="hypothetical protein"
/protein_id="AAL72474.1"
/db_xref="GI:18462702"
/translation="MNHYPSTVTSLETPEARCSGVPPLPACRQRESIYGLIELFTQIV
HRLVNSERLTKTLADQRFVHGKTALLFRIAELANPNPDGLVREVVYHLHIVEPDR
SGKRSSSYLAQLRVDSARGDAVKNRGTLEQDSGLPALYDFGLPRMISTVL"
1176. .2042
/gene="ospB"
/notes="CP0003"
1176. .2042
/gene="ospB"
/notes="similar to a 288 aa protein from Shigella flexneri
emb: CAC05770.1"
/codon_start-1
/transl_table-11
/product="OspB"
/protein_id="AAL72323.1"
/db_xref="GI:18462551"
/translation="MNLDSVRYCRIVNKKNSISDIAFAHIKRVKNSCTHPKAAAL
VFLGKDCSDNDVLSIMGQOIPRYFKNKMLDYFKNEKSKNDFLKMAESMLPOSEP
IIVNDDQDALNAAYEVSVEKAKITVNDDFEYANKYVILHGCSGSHQLGSELID
VOTILSRMKDCGILNVKDIRFTSCGSADKAVKAFNNAPAESLCSILNSIPFFKEKES
LLEQINKHLENDSESDGLKISGYHIGVYHQELFPYSHYRSTSIAPADPEHTVKKRS
QKKTFLINKELD"
2371. .3111
/gene="apy"
/notes="CP0004"
2371. .3111
/gene="apy"
/notes="similar to a 246 aa protein from Shigella flexneri
gb: AAA21206.1"
/codon_start-1
/transl_table-11
/product="apyrase"
/protein_id="AAL72358.1"
/db_xref="GI:18462586"
/translation="MRTKNFLFLCIATNMIFTPSANAKAEGFLTQOTSPDPSILILP
PPAENSVFQADKAHYEFGRSLRDANRVRLASDAYENFGLAFSDAYGMDISRETPR

```

Query Match	27.88;	Score 328.6;	DB 1;	Length 221618;
Best Local Similarity	59.3%;	Pred. No. 1.4e-94;		
Matches 578;	Conservative 0;	Mismatches 394;	Indels 3;	Gaps 1;
QY	210	TGACTAGCAGAGACACAATGATCTGTCGCCGTTCTGTTAATATCAAAACCGGTACTCAATA	269	
Db	193635	TGCTTCGCYAAATGTGAGTATTTTAAAGCAGTTTTATTAATATCAAAACCGCTCTTCTATA	193576	
QY	270	TCTTCTCTGGCGGTGGTGGCCATCATCCGGAAGCGTTCCGGTCCGGGATAAAAAATCGCGC	329	
Db	193575	TCTGTGTTTTGCCTTGGCGAGCCATCTGATAAAAACGCCAGCGGTCAATTAATAAAATCATTC	193516	
QY	330	AGTGC CGGTCTCATGCAGACACATCCCCACGGGTAAACAGCGTCCTGTGCACATCTTC	389	
Db	193515	AGGGCGCACAATCCATCGGTGCTTATTACCTGCTGGTAAACAATGTCCCGGTCGGTTATTC	193456	
QY	390	TGAATGACATCAGGGATCCCGCCGGTCTCACTGGCGGATAACGGGCACGCCGGAGACTGAC	449	
Db	193455	TGAATAACATCAGGGATTCACCAATCTGACTGGCCAGTACAGGCACAGAAAATGCCGAT	193396	

Qy	450	GCTTTCAGCGATACCATACCAAAAGCTTTCATTTTCCGAAGGATGACCAACACACTGGCA	509
Db	193395	GCTTCTGCGCAGTACCATACCAAAAGATTCGTTTTCTGAAGGCAGAAACACCAAGACTGGCA	193336
Qy	510	ATCCGGGTAGACCGGTAAAGCTGGGAAAAGGCACCTGCCATTAACACATCTCCGCTCAT	569
Db	193335	ACCCGATATACGGGGCGGAGGAAAACATATGTGAGCAATAAAACATCATCATGCAATG	193276
Qy	570	CCCAGGTGTTCTGTCTGTGACGACGAGTGTCTGTAATCTTTCACGCGCCGGCGCCAC	629
Db	193275	CCCATACTGTCAATCTGATACTGTAAATGCTCCCGCAGTTCAGGGGAACGACGCCAC	193216
Qy	630	ACGAGCCAGCGAAATGATTTCCCTTCATCTTCAGCTGATACATACACGACGACATAAT	689
Db	193215	ATCAACCAACAGAACTGCTGCTCATTCATTTTAAATGGAACAGTAAATTTCAACATAAT	193156
Qy	690	TCATGTCCTTTTTTCGGGACGTAGATCCCACTCCACGATGAACGATAAGGGAACATTTG	749
Db	193155	TCGTGCCCTTTTTTCGGGCGCAACATTCCTACCTGAGCAATAACGGGGCATCCCTCTCG	193096
Qy	750	GATCAGCCAGGCGGTGATATGACGGGTAAACGGTGCATGGCTTTCATATGCAATGCG	809
Db	193095	GACCCAGCCAAAGAAAGAACATTTGGAGGCAAACTGTTTCGGGAATCAACGTATAA	193036
Qy	810	GGCCAGTCGAACCCGGGTGGAATAACCGTTTACCGGTGTCTGACACCTTCGCCCATC	869
Db	193035	TGGAATCAAAAGCCTGGAGGCACAACAGTAACCCGGGTTCGACATCCTTCTGCTCC	192976
Qy	870	TGCCCCATCATGGGTGAGATAGGCACACAAATGAATACACAGATATTTCAGGGA	929
Db	192975	TGTGCTTCATACTTGTTCCGGGAACAATCACTCATCGCAAAATGATTTATTGAAA	192916
Qy	930	GTTCTGTGCTTACGGGTGATGTAGGTTTTTGTCTGACAATAGTGAAGCGGTG---	986
Db	192915	ACTTTTGTTTTTTCGTGCAATATGTCTTTTGCCGTGATATTCGAAATGGATGTTCT	192856
Qy	987	CATATCAGACGGCTACGTCCTGCTATATTAATGTCATGGCCACTATGCGAGATG	1046
Db	192855	GTGAATAACCGTACTAAACCAACAATATTTGCTATCATGCCCACTGTGACAAAC	192796
Qy	1047	TCAGGTTTAAATTTCCCGATAAATCCGTGCAAGTCTGAGGATGGGAAGGTGAAGG	1106
Db	192795	TTGGGCTGAAACCAACATGAACATTTCCCGAGTAAATCTCCATGCAGTAGGGAT	192736
Qy	1107	TTCCCTGAAAGGAATAAAGTGACATCATGCGCCTCTTTTCTGGGCTTCGGAGCA	1166
Db	192735	TTTTCGAAATACGGGAATGTGATCAATCCCAATTTTACTGGCTTCAAAAGCAAT	192676
Qy	1167	CTTTTTTCTCTGCAG	1181
Db	192675	CTGTTTTTCCCTGCAG	192661

[illegible]

PUBMED 11292750
 REFERENCE 2 (bases 1 to 221851)
 AUTHORS Venkatesan,M.M., Goldberg,M.B., Rose,D.J., Grotbeck,E.J.,
 Burland,V. and Blattner,F.R.
 TITLE Direct Submission
 JOURNAL Submitted (13-FEB-2001) Laboratory of Genetics, University of
 Wisconsin, 445 Henry Mall, Madison, WI 53706, USA
 FEATURES
 source
 1. .221851
 /organism="Shigella flexneri"
 /db_xref="taxon:623"
 /plasmid="virulence plasmid pWR501"
 complement(191..493)
 /gene="S0001"
 complement(191..493)
 /gene="S0001"
 /note="residues 20 to 100 of 100 are 88.23 pct identical
 to residues 217 to 301 of product encoded by
 GenBank Accession Number AE000291 IS2 hypothetical protein
 [Escherichia coli]"
 /codon_start=1
 /transl_table=11
 /product="IS2 orf2, fragment"
 /protein_id="AAK18312.1"
 /db_xref="GI:13310488"
 /translation="MVAHTELMKHAHSPQWDFVEPKNTAVRSPESNRIRAKSFVKTIK
 CDYISIMPKDGLTAARKLAFAFEHYNEWHPHSALDYRSPREYLRQANDNRCLF"
 485..949
 /gene="S0002"
 485..949
 /gene="S0002"
 /note="residues 13 to 93 of 154 are 40.74 pct identical to
 residues 5 to 85 of 333 of product encoded by GenBank
 Accession Number U65001 putative resolvase [Sphingomonas
 sp.]"
 /codon_start=1
 /transl_table=11
 /product="putative resolvase, fragment"
 /protein_id="AAK18313.1"
 /db_xref="GI:13310489"
 /translation="MNHPSVTSLETPEARCRSGVPLPACRORESIYGLIELFIQIV
 HRLSVRSERLVLTLLADFORGHGKTALLPRIAFAALNPDGLVKEVYVHLHVEPDR
 SGRSSSYLAQLRDVSNRGDAVNGRTLPEQDSGLPALYSDPGLPRMISTVL"
 1176..2042
 /gene="S0003"
 1176..2042
 /gene="S0003"
 /note="residues 130 to 273 of 288 are 28.57 pct identical
 to residues 12 to 139 of 340 of product encoded by GenBank
 Accession Number A1139075 ketoI-acid reductoisomerase
 [Campylobacter jejuni]"
 /codon_start=1
 /transl_table=11
 /product="orf, hypothetical"
 /protein_id="AAK18314.1"
 /db_xref="GI:13310490"
 /translation="MNLGVRVPYCRIVNKKNESISDIAFAHIIKRVKNSCTHPKRAAL
 VFLEKGFCDNSVLISGQOIPRFVFNKMLDYVFNKESKNDFLKMAESWLPQSEP
 IVINDDALNAAATSVSKAKIKTVDNDFKEYNKVYIILGHSGPGSHOGLGSELID
 VQIISPMKDCGILNKVDIRFTSCGSADKVPKFNNAEISLCILNISPEPKES
 LLQGIKHLNDESLDGLKISGYHGYVHQQLFPYSHYKSTSI PADPEHTVKKRS
 QKKTFFIINKEL"
 2371..3111
 /gene="S0004"
 2371..3111
 /gene="S0004"
 /note="residues 1 to 246 of 246 are 99.18 pct identical to
 residues 1 to 246 of product encoded by GenBank
 Accession Number U04539 ATP diphosphohydrolase, apyrase
 [Shigella flexneri]"
 /codon_start=1
 /transl_table=11
 /product="apyrase"

/protein_id="AAK18315.1"
 /db_xref="GI:13310491"
 /translation="MKTNLFILFIATNMTIPISANALKARGFLTQOTSPDLSILPP
 PPAEDSVFLADKAHVEGRSLRDANRVRLASEDAYENFGLASDAYGMDISRENTP
 ILQLLTQVLQSDHYAYRNKAKEYKRVRPVVIYKDATCTPDKDEKMAITGTSYSGHA
 SFGWAVALLIAEINQRAEILRGFERGESRVICGAHQWSDVEAGRLMGASVVAVLH
 NTPFTKLSLEAKFEELNPTNELTP"
 3272..3526
 /gene="S0005"
 3272..3526
 /gene="S0005"
 /note="residues 27 to 59 of 84 are 36.36 pct identical to
 residues 427 to 459 of 729 of product encoded by GenBank
 Accession Number D63999 hypothetical protein
 [Synecocystis sp.]"
 /codon_start=1
 /transl_table=11
 /product="orf, hypothetical"
 /protein_id="AAK18316.1"
 /db_xref="GI:13310492"
 /translation="MLPSETMIWOPEFTDKTSRKLGAVPFTTCNVVQGNGLPIPVY
 DOYNDNFRFRAOPKYILGHLSNRLPDTAFENKINH"
 complement(3470..4441)
 /gene="S0006"
 complement(3470..4441)
 /gene="S0006"
 /note="residues 8 to 285 of 323 are 22.56 pct identical to
 residues 56 to 301 of 612 of product encoded by GenBank
 Accession Number AL008970), Hypothetical protein, len: 613
 aa [Plasmodium falciparum]"
 /codon_start=1
 /transl_table=11
 /product="orf, hypothetical"
 /protein_id="AAK18317.1"
 /db_xref="GI:13310493"
 /translation="MIKSLIKYKHRIHTLILKKSLTTILPFGVEISNHOEKLPLN
 KTHTVYFGANAYIIDHSDPYGYMTLTFEHDNAIPPFYHEHQVFLDNFKEVVDVDS
 RYVHGOGKTDVPIENTDMRIGLIGLHIDFISKSDQGFREFCYNNKINDVSLDRII
 NFVQLEYHIPMLSTDNFKIKLRDILEDAIKASNYEENNNKVTDKKMAHQALAYS
 LGDKADIALYLLSKNFETKQDAEMERKNNIYCNLYDVEYLLSKDGCANVKVLEYFI
 NGLVDVNNKFKQKNSGDTMLDNAMKSKDKMIDFFIKKWSGIRQT1"
 complement(4344..4742)
 /gene="S0007"
 complement(4344..4742)
 /gene="S0007"
 /note="residues 38 to 128 of 132 are 23.15 pct identical
 to residues 41 to 123 of 839 of product encoded by GenBank
 Accession Number Z99115 yonO [Bacillus subtilis]"
 /codon_start=1
 /transl_table=11
 /product="orf, hypothetical"
 /protein_id="AAK18318.1"
 /db_xref="GI:13310494"
 /translation="MKNFLKRSIAQSYKMFSGTSPKSLNLSLEAPSGARSFRSL
 EHLDRVSRHYISEIIQKVHPLSSDERHLSIIISNFRHQSNLSNLSNLSNLSKFSFD
 KIQSNITQHNKNTYSEDIKEISNHDVFWG"
 complement(5261..5812)
 /gene="S0008"
 complement(5261..5812)
 /gene="S0008"
 /note="residues 5 to 177 of 183 are 94.21 pct identical to
 residues 42 to 214 of 523 of product encoded by GenBank
 Accession Number AB024946 orf51 [Escherichia coli]"
 /codon_start=1
 /transl_table=11
 /product="ISEC8 orf, fragment"
 /protein_id="AAK18319.1"
 /db_xref="GI:13310495"
 /translation="MSRPSSEINRLKALVAKLQRMQFGKSSSEKLRKTERWQAEQERI
 SALQEMAEITGEQYDVPVLPSSLROSSARKPLPASLPAPRVIREPEECCPACGELS
 PLGCDVSEQLELISSAFVIEKORPKLACRCDHIVQAPVPSKPIARSYAGAGLIAH
 VTGKYADHILPLYRQSDLLFHTAI"
 complement(5742..6164)

CDS
/gene="S0009"
complement(5742..6164)
/gene="S0009"
/note="residues 1 to 116 of 140 are 90.51 pct identical to
residues 1 to 116 of 272 of product encoded by GenBank
Accession Number X05952 pot. ORF B [Shigella sonnei]"
/codon_start=1
/transl_table=11
/product="IS3 orfB, fragment"
/protein_id="AAK18320.1"
/db_xref="GI:13310496"
/translation="MCRVFGVSRSYDVRQVHAPSQRKOSDERLKLKIVAHIRRET
YFRRIGTELADNGITVGRDLRLRLKELHCKOKRFRATNSDHLNLPVFNLLNQ
NTPFPAPNWDVSVQAFRNPQTEGARETAAYR"
complement(6200..6502)
/gene="S0010"
complement(6200..6502)
/gene="S0010"
/note="residues 1 to 100 of 100 are 93.99 pct identical to
residues 1 to 100 of 100 of product encoded by GenBank
Accession Number X05952 pot. ORF A [Shigella sonnei]"
/codon_start=1
/transl_table=11
/product="ISEc8 orf"
/protein_id="AAK18321.1"
/db_xref="GI:13310497"
/translation="MSKNQRYSTEFKAEAVKTVVPENQLISISGASRLSLPEGTLOW
VTAARKGLNTPGSRVTAELSEYVQLRKALNEARLERDILKATAYFAFBSLKNR"
6793..7110
/gene="S0011"
6793..7110
/gene="S0011"
/note="residues 1 to 58 of 105 are 34.48 pct identical to
residues 383 to 437 of 1003 of product encoded by GenBank
Accession Number AL023835 similar to Src homology domain
3; cDNA EST yk491e6.5 comes from this gene; cDNA EST
EMBL:D36655 comes from this gene; cDNA EST yk200b8.5 comes
from this gene; Score 328.6; DB 1; Length 221851;
Best Local Similarity 59.3%; Pred. No. 1.4e-94;
Matches 578; Conservative 0; Mismatches 394; Indels 3; Gaps 1;
QY 210 TGACTAGCCAGACAAATGATCTGTGCGGTCTGTATATCAAAACGGTACTCAATA 269
DB 184983 TGGTTCGTAATGTAGTATTTTAAACGAGTTTATATATCAACCGCTCTCTATA 184924
QY 270 TCTTCTGTGCGTGGTGCATATCCGGAAGCGTTCCGGTCGGGATAAATAAATCGCG 329
DB 184923 TCCTGTTTTCCTGGCGAGCCATCTGATAAACGCCAGGGTCAATTAATAAATCATTC 184864
QY 330 AGTGGCGCGTCCATGACAGACATCCCGCCAGGTAACAGCGTCCCTGTACATTCCTC 389
DB 184863 AGGGCGCACATCCATCGTCTTATACCTGCTGGTAACAAATGTCGCGGTCTGGTTATTC 184804
QY 390 TGAATGACATCAGGATCCCGCGCTCACTGCGATACGCGGACGCGGAGACTGAC 449
DB 184803 TGAATACATCAGGATTCACCAATCTGACTGGCAGTACAGGACAGAAATGCGGAT 184744
QY 450 GCTTCAGCCAGTACCATACCAACCGTTTCAATTTCCGAGGAGTACACACACACTGGCA 509
DB 184743 GCTTCAGCAGTACCATACCAACGATTCGTTTCTGAGGAGAGACACACACTGGCA 184684
QY 510 ATCCGGTAGACGGTAGCTGGGAAAGGCGACCTGCCATTAACACATCTCCGCTCAT 569
DB 184683 ACCCGATATACGGGGCGGAGGAAACAACTGTCAGCAATAAAACATCATCATGATG 184624
QY 570 CCCAGGTGTTCTGCTGACGACGAGTGTCTGCTATTTCTACCGCCGCGGCCACC 629
DB 184623 CCCATGCTGCAATCTGATACTGTAATGTCCTCCGAGTTCAGGCGAACACAGCCACCA 184564
QY 630 ACCAGCGAGGGAATGATTTCCCTTCATCTCCAGTGTATACATACACGACGACATAAT 689
DB 184563 ATCAACCAACAGACTGCTGCTCCATTCATTTTAAATGGACAGTAATTCACATTAAT 184504

QY 690 TCAATGCTCTTTTTCGGGACGTAGCATCCCACTGAAACAGCAATGATGTGCT 749
DB 184503 TCGTGCCCTTTTTCGGGCGCAACATTCCTACCTGAGCAATAACGGGCAATCCCTTCGG 184444
QY 750 GATGACGCCAGCGGTGGATATCGAGGGTAAACGGTTCGATGCTTCATATCAATGCG 809
DB 184443 GAGCCAGCCAGAAAGAACATTTGGAGGCAAACTGTTGGGGAATCAACGTATAATTC 184384
QY 810 GGCAGTCGAAACCCGGTGAATAACCGTTTACCGGTCTCTGACACCTCCGCCATCAGA 869
DB 184383 TGGAAATCAAGCCTGGAGGCACACAGTAACCCGGGTTGACATCCTTCTGCTCCAGA 184324
QY 870 TGGCCCATCATGGGTGAGATAGGCAACAATAAATAACACAGATAATTCAGGGAAC 929
DB 184323 TGTGTCTTCATCTGTTTCCGGGAACAATCACCTCATCGCAAAATGATTTATTGAAAA 184264
QY 930 GTTCTGCTTACCGGTGATGATAGTTTGTCTGACATAGTGAAGCGTG--ACAG 986
DB 184263 ACTTTGTTTTCGTGCAATATGCTTTTGCTGATAATCTGAATGATGTTTCCTA 184204
QY 987 CATATCAGAGCGTCACTGCTGCTATATTAATCTGATGGCCACTATGGCAGATGACACAGA 1046
DB 184203 GTGAATAACCGTACTAAACCAACAATATGCTATCATGCCCCTGTGACAAACGATTGCA 184144
QY 1047 TCAGGTTTAAATTCGCCGATAATCCGTCGAAGTCTGAGGATGGAAGGAGTGAAGCTG 1106
DB 184143 TTGGGTGAAACCATGAATATTCGAGTAATCTCCATGATGAGGATGTGAAGACTG 184084
QY 1107 TTCTTGAAGGAATAAAGTGACATCATGCCCTCTTTTCTGCTCCGAGCAATTTTA 1166
DB 184083 TTTGGAATACCGGATGATATCAATCCCAATTTACTGGCTTCTTCAAAAGCAATTTTG 184024
QY 1167 CTTTTCCTGTCGAG 1181
DB 184023 CTGTTTTCCTGTCGAG 184009

RESULT 12
AF134403/c
LOCUS
DEFINITION
3500 bp DNA linear BCT 02-JUN-1999
Escherichia coli plasmid pAA2 Shf (shf), hexosyltransferase homolog
(capD), and VirK (virK) genes, complete cds.
ACCESSION
AF134403
VERSION
AF134403.1 GI:4959583
KEYWORDS
SOURCE
ORGANISM
Escherichia coli.
Escherichia coli.
Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
Escherichia.
REFERENCE
1 (bases 1 to 3500)
AUTHORS
Czeczulin,J.R., Whittam,T.S., Henderson,I.R., Navarro-Garcia,F. and
Nataro,J.P.
Phylogenetic analysis of enteroaggregative and diffusely adherent
Escherichia coli.
INFECTION
Immun. 67 (6), 2692-2699 (1999)
JOURNAL
MEDLINE
PUBMED
10338471
REFERENCE
2 (bases 1 to 3500)
AUTHORS
Czeczulin,J.R., Whittam,T.R., Henderson,I.R. and Nataro,J.P.
Direct Submission
TITLE
Submitted (11-MAR-1999) Vaccine Development, Univ. of Maryland, 685
W. Baltimore St., Baltimore, MD 21201, USA
FEATURES
Location/Qualifiers
source
1..3500
/organism="Escherichia coli"
/strain="042"
/db_xref="taxon:562"
/plasmid="pAA2"
/note="enteroaggregative"
558..1400
/gene="shf"
558..1400
gene
CDS

Query Match	20.9%	Score	246.6	DB 1	Length	3500			
Best Local Similarity	56.4%	Pred. No.	3e-68						
Matches	550	Conservative	0	Mismatches	379	Indels	46	Gaps	3

QY	210	TGACTAGCCAGGACACAATGATCTGTGCGGTTCTGTTAATATCAAAACGGGTACTCAATA	269
DB	2441	TGCTTCGGTAAATGAGTATTTTAAACACAGTTTTATTAATATCAACCGCTCTCTATA	2382
QY	270	TCTTCTTGCGGCTGGCTGCCATCATCCGGAAGCGTTCCGGTGGGATATAAAATTCGGGC	329
DB	2381	TCTCTGTTTTCGACGGCGAGGCATCTGATAAAACGCCAGGGTCAATATAAAATCATTC	2322
QY	330	AGTGGCGGCTCCATCGAGACACATCCCCACGGGTAAACAGCGTCCCTGTCACTTCCTTC	389
DB	2321	AGGGCGCACATCCATCCGTGGTGTATTAACCTGCTGGTAAACATGTCCGGTCTGGTTGTC	2262
QY	390	TGAATGACATCAGGATCCCGCCGTCCTCACTGGCGATAACGGGCACGCCGAGACTGAC	449
DB	2261	TGAATAACCTCAGGGATTCCACCAATCTGATGGCCACTACAGGCACAGAAATGCCGAT	2202
QY	450	GCTTCAGCCAGTACCATACCAACCGCTTCATTTTTCGAAAGCATGACCACCACACTGGCA	509
DB	2201	GCTTCTGGCAGCATTACCAAAAGATTCGTTTTCTGAAGGCAGAACCAACCAAGACTGGCA	2142
QY	510	ATCCGGTAGACCGGTTAAGCGTGGGAAAAGGGCACCTGCCATTAAACACATCTCCGCTCAT	569

Japan
Phone: 03-3443-8111x252
Fax: 03-3443-3893
Location/Qualifiers
1. .1642

FEATURES

source
gene
CDS
/organism="Shigella flexneri"
/db_xref="taxon:623"
/plasmid="Plasmid pMYSH6000"
373. .1323
/gene="virk"
373. .1323
/gene="virk"
/codon_start=1
/transl_table=11
/product="virulence protein"
/protein_id="BAA01778.1"
/db_xref="GI:216944"

translations="MFSVSNLSIFGLKRVFSSDLSLPGKWEHRKFRFMYILRCALNP
VASIRVYELRSIQCTEDILAIQPTLPARITRPLHLKGGRAWRSQVILRHYFVONL
PEKYSBELPKSVSLVQFIGKDEDFDIQCSFGDFDEGELMSLFFNKKIVIRATLF
SVILITQNGHTAFIGLQAGAPKNTGPDILIRCATRACYGLFPKRIIFAFCAKMACNVS
ECIAVSEHSVIFRLRYWYOKRKTFFVAVSDFEWAGTKCGDWYKLPQTQVVRKPLSN
IASKKRSEYRKRYALDLYIHETAIRSLDAPVNSEHQDLN"

BASE COUNT 451 a 311 c 376 g 504 t
ORIGIN

Query Match 13.5%; Score 159.4; DB 1; Length 1642;
Best Local Similarity 65.1%; Pred. No. 6.7e-40;
Matches 235; Conservative 0; Mismatches 126; Indels 0; Gaps 0;

```

QY 210 TGAAGCCAGACACAAATGATCTGTGCGGTTCTGTTAATATCAAAACCGGTACTCAATA 269
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 361 TGCTTCGCTAATGTAGTATTTTAAACGAGTATTTAATATATCAAAACCGGTCTTCTATA 302

QY 270 TCTTCTCGCGTGGCTGCATCATCCGGAAGGTTCCGGTGGGATAAAATACCGC 329
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 301 TCCTGTTTCCCTGGCGAGCCATCTGATAAAACGCCAGGTCATTAAAAAATCATTC 242

QY 330 AGTGGCGCGGTCCATGACACACATCCCGGAGTAAACAGCGTCCCTGTACATTTCTC 389
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 241 AGGCGCACATCCATCGTGTATTTACCTGTGTAACAAATGTCGCGGTCTGTTTATTC 182

QY 390 TGAATGACATCAGGATCCCGCGCTCTCACTGCGGATACGGCAGCGCGGAGACTGAC 449
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 181 TGAATACATCAGGATTCACCAATCTGATGCGCAGTACAGGCACAGAAAATGCCGAT 122

QY 450 GCTTCAGCCAGTACCAACAGCTTCATTTCCGAAGGATGACACACACTGGCA 509
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 121 GCTTCGCCAGTACCAACAGGATTCGTTTCTGAAGCAGAACCCAGACTGGCA 62

QY 510 ATCCGGTAGACCGGTACGCTGGGAAAGGACCTGCCATTACACATCTCCGCTCAT 569
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 61 ACCGATATACGGGGCGGAGGAAACATTTGTCAGCAATAAAACATCATCATGATG 2

QY 570 C 570
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1 C 1

```

RESULT 14

E06635/c
LOCUS E06635 1642 bp DNA linear PAT 29-SEP-1997
DEFINITION Shigella virK gene.
ACCESSION E06635
VERSION E06635.1 GI:2174818
KEYWORDS JP 1994030766-A/1.
SOURCE unidentified.
ORGANISM unidentified.
1 (bases 1 to 1642)
Yoshikawa,M.
ATTENUATED LIVE VACCINE OF DYSENTERY BACILLUS HAVING HIGH SAFETY
Patent: JP 1994030766-A 1 08-FEB-1994;
JOURNAL

RES DEV CORP OF JAPAN
OS Shigella flexneri 2a
PN JP 1994030766-A/1
PD 08-FEB-1994
PF 15-JUL-1992 JP 1992210772
PI YOSHIKAWA, MASANOSUKE
PC C12N1/21,A61K39/00,A61K39/00,A61K39/112,A61K39/112,C12N15/31,
C12P21/02,
PC C12N1/21,C12R1:01),(C12P21/02,C12R1:01);
CC strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
CC *source: strain=YSH6000T;
CC *source: clone=pMYSH6000;
FH Key Location/Qualifiers
FT mat_peptide 373. .1320
FT Location/Qualifiers
FT 1. .1642
FT /organism="unidentified"
FT /db_xref="taxon:32644"

FEATURES

source
BASE COUNT 451 a 311 c 376 g 504 t
ORIGIN

Query Match 13.5%; Score 159.4; DB 6; Length 1642;
Best Local Similarity 65.1%; Pred. No. 6.7e-40;
Matches 235; Conservative 0; Mismatches 126; Indels 0; Gaps 0;

```

QY 210 TGAAGCCAGACACAAATGATCTGTGCGGTTCTGTTAATATCAAAACCGGTACTCAATA 269
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 361 TGCTTCGCTAATGTAGTATTTTAAACGAGTATTTAATATATCAAAACCGGTCTTCTATA 302

QY 270 TCTTCTCGCGTGGCTGCATCATCCGGAAGGTTCCGGTGGGATAAAATACCGC 329
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 301 TCCTGTTTCCCTGGCGAGCCATCTGATAAAACGCCAGGTCATTAAAAAATCATTC 242

QY 330 AGTGGCGCGGTCCATGACACACATCCCGGAGTAAACAGCGTCCCTGTACATTTCTC 389
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 241 AGGCGCACATCCATCGTGTATTTACCTGTGTAACAAATGTCGCGGTCTGTTTATTC 182

QY 390 TGAATGACATCAGGATCCCGCGCTCTCACTGCGGATACGGCAGCGCGGAGACTGAC 449
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 181 TGAATACATCAGGATTCACCAATCTGATGCGCAGTACAGGCACAGAAAATGCCGAT 122

QY 450 GCTTCAGCCAGTACCAACAGCTTCATTTCCGAAGGATGACACACACTGGCA 509
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 121 GCTTCGCCAGTACCAACAGGATTCGTTTCTGAAGCAGAACCCAGACTGGCA 62

QY 510 ATCCGGTAGACCGGTACGCTGGGAAAGGACCTGCCATTACACATCTCCGCTCAT 569
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 61 ACCGATATACGGGGCGGAGGAAACATTTGTCAGCAATAAAACATCATCATGATG 2

QY 570 C 570
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1 C 1

```

RESULT 15

AF326777
LOCUS AF326777 66714 bp DNA linear BCT 18-JUN-2002
DEFINITION Shigella flexneri 2a SRL pathogenicity island, complete sequence.
ACCESSION AF326777
VERSION AF326777.2 GI:21450881
KEYWORDS
SOURCE Shigella flexneri 2a.
ORGANISM Shigella flexneri 2a.
Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
Shigella.
1 (bases 1 to 66714)
REFERENCE Luck,S.N., Turner,S.A., Rajakumar,K., Sakellaris,H. and Adler,B.
AUTHORS ferric dicitrate transport system (Pec) of Shigella flexneri 2a
TITLE

YH6000 is encoded on a novel pathogenicity island carrying multiple antibiotic resistance genes
 Infect. Immun. 69 (10), 6012-6021 (2001)
 21437601
 PUBMED 11553538
 REFERENCE 2 (bases 1 to 66714)
 LUCK, S.N., TURNER, S.A. and RAJAKUMAR, K.
 Direct Submission
 Submitted (06-DEC-2000) Microbiology, Monash University, VIC 3800, Australia
 REFERENCE 3 (bases 1 to 66714)
 LUCK, S.N., TURNER, S.A. and RAJAKUMAR, K.
 Direct Submission
 Submitted (17-MAY-2002) Microbiology, Monash University, VIC 3800, Australia
 REMARK Amino acid sequence updated by submitter
 REFERENCE 4 (bases 1 to 66714)
 LUCK, S.N., TURNER, S.A. and RAJAKUMAR, K.
 Direct Submission
 Submitted (18-JUN-2002) Microbiology, Monash University, VIC 3800, Australia
 REMARK Sequence update by submitter
 COMMENT On Jun 18, 2002 this sequence version replaced gi:15808696.
 FEATURES
 Location/Qualifiers
 1..66714
 /organism="Shigella flexneri 2a"
 /db_xref="taxon:42897"
 /note="corresponds to Escherichia coli min 23"
 458..66714
 /note="SRL pathogenicity island"
 619..1836
 /gene="int"
 619..1836
 /gene="int"
 /note="member of the P4 bacteriophage integrase family"
 /codon_start=1
 /transl_table=11
 /product="integrase"
 /protein_id="AAL08426.1"
 /db_xref="GI:15808697"

misc_feature

gene

CDS

CDS

CDS

CDS

CDS

CDS

repeat_region

repeat_region

CDS

CDS

/protein_id="AAL08428.1"
 /db_xref="GI:15808699"
 /translation="MHNNMNSNOTGQYTPSRILMAELRLLGILSRSTIVKINPA
 SKYHPTPKPIRLGSAVGVWRASAIWELMLHTAPAWSEPNKELNDDDKITP"
 3868..5730
 /note="Orf4"
 /codon_start=1
 /evidence=not_experimental
 /transl_table=11
 /product="unknown"
 /protein_id="AAL08429.1"
 /db_xref="GI:15808700"
 /translation="MOTNRQAQMTFTNMLYDICTVLGFKGENPAHKRKSNTFEM
 HKFWDQRYNEISRIIDAEQVFSQQRRIIYARIEHYFYMNSYPVHSTLKEPELRSY
 CLRTFGLVVDYNTYRPENDSAFYIHYINFLQKSYCPCLDHADTSDAAVKRYL
 REYLAELGFKNDPHENKVALGKRYTGTIRKDNKSKSLMOQYIMAKNEYKKDYRE
 KKLDELEKFLNRIDKFYNAPYSLSVLDDIQRTKILQSLAYLVLYRGLWHGL
 YGAAQYLYDFTSFTTTPYAKLLEMFYKQNSAEGTSLRSYSLDDSKQYIIRLKL
 LVFNINDKNGCDDAYLKKIISYFQQLQNEAVHVTSCTETLAVYICLIIRKNKINDVQ
 YDDMERKGLFGLPSPGVYRGALSLLRTALEVKNRKNIKYSLFYWLDHVAYQDAFI
 EKIPLDIVYKEGEIQYDANNFTLMRVIKYKNCLEKISTKPTAPYITGLDDVEK
 VLDKINILDKYEVYDGTALAEVIMENKVLSSRERKMTIGLTGSKKYTLQCVKL
 GVLVHVKSVPDEIKNVMLYGDKAENRRRMIDALTICEDDIRNPPFLS"
 7104..7697
 /note="Orf5"
 /codon_start=1
 /evidence=not_experimental
 /transl_table=11
 /product="unknown"
 /protein_id="AAL08430.1"
 /db_xref="GI:15808701"
 /translation="MYNANPNYKMDFMILKDVNEHMDGMFORFKLLPFRIDFAYRKD
 YPFGSHCHSMCEIYRLSETQTMAGYVWMEYTPDKGLHIFHIGYLDGQRHNS
 YQISQGLGDIWRITTEGEGYFHLCKRAKDKYVPRIDHVIHYSKSAVDLRLVALSYLAK
 ODOKEGIIIGRSRLPEKSNRGRPDITELWGLMSELS"
 8449..8745
 /note="Orf6; similar to yfjJ on the cryptic P4 prophage
 CP4-57 of Escherichia coli K-12"
 /codon_start=1
 /evidence=not_experimental
 /transl_table=11
 /product="unknown"
 /protein_id="AAL08431.1"
 /db_xref="GI:15808702"
 /translation="MSLSRLFFFNKDAYHYHGNYSRDTLRLGLITGAWYSALGLELED
 HGQLVHFPAAGRYILDINDPNFEQYQRLVRLDYTLTKLDTKASQQRNFYSGF"
 complement(9112..10421)
 /note="IS629-like element"
 10422..>10905
 /note="remnant of IS200-like element"
 complement(10980..11918)
 /note="Orf7; similar to a LysR-like transcriptional
 regulator"
 /codon_start=1
 /transl_table=11
 /product="putative transcriptional regulator"
 /protein_id="AAL08432.2"
 /db_xref="GI:21450882"
 /translation="MKNITKWLXDFLTLEACRHFSSQAAKERNLSQAFSRRIKALEA
 AIGVLFRTTTPLTQTEGKLFHSQTSRLQLQECNLGELNQLSGVLPNKLAAAH
 SLSLVPRLVHSLTAYGGEFYHVEADIVQAVNTLRGSKDFIISPRDLMQSPF
 CCLIKFSELSLPVCAADQAQHPVDITQPOVPLNYTATSYMGLVNRHRLAEVGSITG
 RTIFSSMSELKNNALNGYGIAMLPINISIVDELQTKRLICLDAAKLTVPVQAYIRM
 NTRLNRTAENWRIQLQEHMPDDLIQIISMEPPARN"
 12299..12988
 /note="Orf8"
 /codon_start=1
 /evidence=not_experimental
 /transl_table=11
 /product="unknown"
 /protein_id="AAL08433.1"
 /db_xref="GI:15808704"
 /translation="MGPATVDLFPNKFTVTTAAOCDQEHILPILISSIPDIPDRTDALM


```

VNHRYGKPKKATEDDKSMEEIEGDLVFLVVLCLANSIDISLEAHDRVMHKNFNRDKD
RWTKEGK
1358. .2101
/gene="dapB"
1358. .2101
/gene="dapB"
/function="lysine biosynthesis"
/note="28.2% of identical aminoacids with the Escherichia
coli dihydropicolinate reductase; putative"
/codon_start=1
/transl_table=11
/product="dihydropicolinate reductase"
/protein_id="AAA92874.1"
/db_xref="GI:755603"
/transl_table="MSNETIKLVIAGPRGMQAEVAKLERTPHFDLVGAIDHTYDQ
KLSDVVEVSADFTYDIDHACFTPTQDVLIDLTPEIGKVHTKIALEHGVVRPVGTT
GFSEADLKEITTEEGIGAIIPNFAIGALIMKMFMAANYFEDVEIIEHDDOK
LDAPSGTALTAEMI SEVKEKQGHDEKILLPGARGAQNGIRLRLHRIQIRHDSYNR
ASPSGVKLSVEQVMKIDQLVGLNLIID"
2117. .2530
/gene="joJF"
2117. .2530
/gene="joJF"
/note="putative"
/codon_start=1
/transl_table=11
/product="unknown"
/protein_id="AAA92875.1"
/db_xref="GI:755604"
/transl_table="MKALIAHDKKQDMVQFTAYRDILKNHDLVATGTTGLKIEA
TGILQEFQSGPLGGDQOIGALIAANALDLVFLRDLPLAQHPEDVSALIRLDVYS
IPLATNNGRAEILVRLDEGVFEFOLLRGEENV"
2523. .3197
/gene="joJG"
2523. .3197
/gene="joJG"
/note="putative"
/codon_start=1
/transl_table=11
/product="unknown"
/protein_id="AAA92876.1"
/db_xref="GI:755605"
/transl_table="MYNADVLAFGASHDDVEIGMGGTIAKFKQEKVMICDLTAEAL
SSNGTVSLRKEEAARILGADKRIQLILPRLGLMSDAQISVTVIRICRPAVF
MPYKDRPHDGNALVEAIPFSGIHKDKESLPFAHKVSKVYIMINGPHQDFV
IDISDTIAKKQSLNAYSOQIPSKDSVSTPLTNGYIETVEAREKLYGREAGVEYAEV
SFPNGC"
3229. .4302
/gene="joJH"
3229. .4302
/gene="joJH"
/note="21.6% of identity to trans-acting transcription
factor of Saccharomyces cerevisiae; 25% identity to sucrose
synthase of Zea mays; putative"
/codon_start=1
/transl_table=11
/product="unknown"
/protein_id="AAA92877.1"
/db_xref="GI:755606"
/transl_table="MRKLGKLAEGHEIHFTSSIPRLNTHPHNIHFHEVNOYA
VFYPPYDITLASKIAEVAERENLDIHAHYALPHAVCAYLAKOMLKRNGIVTTVLHG
TDITVLGYDPSLDLIRFALESSDRVTAYSSALAAETDYLKPEKTIETVNFIDERV
YLAKNAAIKKEGILPDEKVVHVSFRKVRQDVIRVFRNIAGKAKALLVGDG
PEKSTACELIRKGLQDVLMGNQDRVEDLISDLKLLSEKESFGLVLEAMACG
VPCIGNIGGIPVKNVSGFLVDVGDVTAATARAMSILEDEQLSNRFTKAAIEMLE
NEFSRKIVSQYBQIADLAEP"
4307. .5440
/gene="paps"
4307. .5440
/gene="paps"
/note="38.9% of identical aminoacids with the Escherichia
coli pcnB gene product, encoding poly(A) polymerase;
putative"

```

```

/codon_start=1
/transl_table=11
/product="polyA polymerase"
/protein_id="AAA92878.1"
/db_xref="GI:755607"
/transl_table="MEKVFIRKALPVLIRILBAGHOAYFVGAVRDSYMKRTTIGDVDA
TDAAPDQVERLFTQVDYGDSDYDFRSEVQFISSLEEDLKRDLTINAMAWTA
DGKVLIDYFGKKIDQKIVTVGKPEDRFQDALRMLRAVRFMQGLFTLSPTEEEAI
AKEKSLLSHVSVERTIEFKLQGRASQALOTLIOTRVEELPGVYHRENLIS
EPFPLSTRRELNAALLINGVLKADPLKAWKLPKVIKEAIIHADTFGQSLDA
MTMYRAGKALLSAKISQLRQNEKLEKKLDIQAYQNLPIKSLKDLDTIGKDLIA
LNRNPAKWVSELEQWIEQAVVTGKLSNQKHIEEWLTKTQGH"
5425. .6342
/gene="bira"
5425. .6342
/gene="bira"
/function="bifunctional biotin operon repressor"
/note="26.7% identity to the Escherichia coli bifunctional
biotin operon repressor"
/codon_start=1
/evidence=experimental
/transl_table=11
/product="biotin [acetyl-CoA carboxylase] ligase"
/protein_id="AAA92879.1"
/db_xref="GI:755608"
/transl_table="MRSTLRKDLIELFSQAGNEFISGOKISDALGCSRATVWKHIEEL
RKEGYEVAERKGVRLIKPKLSEIRFGLKTEVMGOHLIYHDLVSSQTKTAHEL
ANNNAECTLEGNGVMSLILRPDIPLOKTPQLTLLAAVAVQGIIEAAGTQTDKWP
NDLLNGKTKYTGILTEMOAEDRVSRTIIGIIVNQPNDFPDELKDIAISLSQAG
EKIDRAGVQHILLCFERKRYDMTHGTFPIKLLWESLALGIGTNMRARTLNGTFYK
ALGIDDESVLLELTNEGKIKIYSADIEIG"
5501. .6533
/note="putative"
termInator
BASE COUNT 2036 a 1318 c 1547 g 1639 t
ORIGIN
Query Match 5.4%; Score 64.2; DB 1; Length 6540;
Best Local Similarity 54.4%; Pred. NO. 9.9e-09;
Matches 129; Conservative 0; Mismatches 108; Indels 0; Gaps 0;
QY 335 GCGGTCCATGCAGACATCCCCACGGGTAAACAGCGTCCCTGTCACATTTCTGNAAT 394
Db 4158 GCCTTTGCGGCGGTAAATCACCACATCCACAAAATCCGCTCACATTTTAT 4099
QY 395 GACATCAGGATCCCGCGTCTCCTGCGGATACGGCGACGGCGGAGACTGACGCTTC 454
Db 4098 AACCTCAGGATACCGCAATGTTTGTTCATACAGGCACCTCGCAAGCATCGCTTC 4039
QY 455 AGCCAGTACCATACCAACGGTTTCATTTCCGAAGGCATGACCACACATGGCAATCCG 514
Db 4038 AAGCAGGACAAAGGCGCAAGCTTTCTTTTCAGATAGCAGCAGCTTCAATCGCTAATAGA 3979
QY 515 GTAGACCGGTAAACGCTGGGAAAGGCGACCTGCCATTAACACATCTCGGCTCATCC 571
Db 3978 ATAAAGATCTTCAACACCGCTCTTGATTTCCAGCATTAAGACTTGGTCTTCCAAGCC 3922
RESULT 17
AX100832
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
AX100832
Sequence 76 from Patent WO0121772.
AX100832
AX100832.1 GI:13619763
synthetic construct.
synthetic construct
artificial sequences
1 (bases 1 to 10801)
Iocum,R.R., Patterson,T.A., Hermann,T. and Pero,J.G.
Methods and microorganisms for production of panto-compounds
Patent: WO 0121772-A 76 29-MAR-2001;
OMNIGENE BIOPRODUCTS, INC. (US)
Location/Qualifiers

```

```
source          1. .10801
                /organism="synthetic construct"
                /db_xref="taxon:32630"
                /note="Recombinant pAN240 plasmid"
BASE COUNT      2797 a 2485 c 2151 g 3368 t
ORIGIN
Query Match      5.4%; Score 64.2; DB 6; Length 10801;
Best Local Similarity 54.4%; Pred. No. 1.1e-08;
Matches 129; Conservative 0; Mismatches 108; Indels 0; Gaps 0;
Qy 335 GCGGGTCCATGCAGACACATCCCCACGCGGTAAACAGCGTCCCTGTGCACATCTTCTGAAT 394
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 6490 GCGTGTGCGCGGTACATCAGCAGATCCACCAAAATCCGCTACATGTTTTTAT 6549
Qy 395 GACATCAGGATCCCGCCCGTCTCAGTGGGATACGCGGACCCGCGAGACTGACGCTTC 454
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 6550 AACCTCAGGATACCGCCCAATGTTTGTTCATACGAAGCACTCCGCAAGCCATCGCTTC 6609
Qy 455 AGCCATACCATACCAACGCTTCATTTCCGGAAGCATGACCAACATCGCAATCCG 514
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 6610 AAGCAGACAAAGCGAAGCTTTCTTTTCAGATAGCAGAGTCAATCGCTAAATAGA 6669
Qy 515 GTAGACGGTAACGCTGGGAAAAGGCGACCTGCATTAAACATCTCCGCTCATTC 571
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 6670 ATAAAGTCTCAACAGGCTTGATTTCCAGCATTAAGACTTGGTCTTCCAAGCC 6726

RESULT 18
LOCUS BACYP1A/c
DEFINITION Bacillus subtilis (clone YAC15-6B) yPIABF genes, qcrABC genes,
          yPjABCDfGH genes, bira gene, panBCD genes, ding gene, ypmB gene,
          aspB gene, asnS gene, dnaB gene, nth gene and ypoC gene, complete
          cds.
ACCESSION L47709
VERSION L47709.1
KEYWORDS GI:1146223
SOURCE Bacillus subtilis.
ORGANISM Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
REFERENCE 1 (bases 1 to 23775)
AUTHORS Henner, D., Gollnick, P. and Moir, A.
TITLE Analysis of an 18 kilobase pair region of the Bacillus subtilis
        chromosome containing the mtr and gerC operons and the aro-trp-aro
        supraproton
JOURNAL Proc. Int. Symp. Genet. Ind. Microorg. 6, 657-665 (1990)
REFERENCE 2 (bases 1 to 23775)
AUTHORS Sorokin, A., Azevedo, V., Zumstein, E., Galleron, N., Ehrlich, S.D. and
        Serror, P.
TITLE Sequence analysis of the Bacillus subtilis chromosome region
        between the .serA and kdg loci cloned in a yeast artificial
        chromosome
JOURNAL Microbiology 142 (Pt 8), 2005-2016 (1996)
MEDLINE 96349105
PUBMED 8760912
FEATURES
    source          1. .23775
                    Location/Qualifiers
                    /organism="Bacillus subtilis"
                    /strain="Marburg"
                    /sub_strain="168"
                    /db_xref="taxon:1423"
                    /map="202 degrees"
                    /clone="YAC15-6B"
                    /notes="bases 57126 to 80902 of the YAC15-6B clone"
                    1. .33
                    /gene="aroE"
                    1. .33
                    /gene="aroE"
                    /standard_name="transcription terminator"
                    /note="putative"
                    /citation=[2]

gene
terminator

348. .1619
/gene="ypIA"
348. .1619
/gene="ypIA"
/function="hypothetical"
/note="putative"
/citation=[2]
/codon_start=1
/transl_table=11
/protein_id="AAB38432.1"
/db_xref="GI:1146224"
/translation="MNTLIOEALKLVKAEGETEKLNTLSAEKQLHDEKDAIAAQLYY
EWGDEKATSLISDLHDLYNETHLNFYAEILLIDDEEKALAVLETTIPETDPSTPE
SLIMADLYOMQGVLFVSEOKLFOAKSILNDPVIDFALGELYFAQAGYAKAVOYFKT
TAEQSEIGVGNVHVRSLAESLSGSEFEDAIPWEKAVDNDPDPNTIFGYPTALQAG
LVKTAIKQLSDLKELDPSTSLYMPLSKSYEABGMYEALKTKAREGIRYDEYNKELFL
YAAKMAIKGKSEEGKLLQELALDLPGFVEALHTLLAVYHKEDYDQIIDLQEVRS
YGEDPKYNNLASAYTELEQYEEAKQSFEAAVLYHVEDRDFLYEYASFLLEEGLORE
ALPLKKVLEMDGANELEETILRIEDEFSR"
1691. .2230
/gene="ypIB"
1691. .2230
/gene="ypIB"
/function="hypothetical"
/note="putative"
/citation=[2]
/codon_start=1
/transl_table=11
/protein_id="AAB38433.1"
/db_xref="GI:1146225"
/translation="MOTPVSVNEKKDFIRWLNHYQKRRCEVWILNLMSHDLSMEK
VHFVQAEFCPRGIIIMSTHCVEVPFRFYKENVNTYTDKRSFHDRLNKQODLFIQLN
FRSAGSYPEYAAVLESNPHIPKPNLFENKKDQGLAEQILEHAISTFOREKLLKDIDDL
DRDKKAEPLSLRQLNLT"
2285. .2731
/gene="ypIF"
2285. .2731
/gene="ypIF"
/function="hypothetical"
/note="putative"
/citation=[2]
/codon_start=1
/transl_table=11
/protein_id="AAB38434.1"
/db_xref="GI:1146226"
/translation="MRWRITDAKDYLOAKDYIDTAVIPLINIRVNNHFKMAAEKGETF
QLLSEELEROLKGRVYLLPPYTVDRNEITVOGLKUREELITEFPHVLLTSDSNR
AEDALCKMIVTSSVPLEHLNDSLKRKILDERTAELNLVLQLNSTL"
2771. .2776
/citation=[2]
/evidence=experimental
2794. .2799
/citation=[2]
/evidence=experimental
2874. .3377
/gene="qcrA"
2874. .3377
/gene="qcrA"
/function="menaquinone oxidase"
/standard_name="Rieske protein"
/note="26.2% of identical aminoacids with the
cyanobacterium Nostoc PCC 7906 Rieske iron-sulfur protein"
/codon_start=1
/evidence=experimental
/transl_table=11
/product="Rieske iron-sulfur protein"
/protein_id="AAB38435.1"
/db_xref="GI:1146227"
/translation="MGKGDHISRROFLNLTGTGGVGMAASMLMPVRFALDPVLKST
KQDMYQVVSVDLETKEPORFDKINQVDWYSESRSAWFKNGDEIVALSPICKH
LGCTVWNSDPKPNKFFCFCHYGLIEKDGNTNVPGLPPLADPHYEQEVKDFLYLQK
AKPKGE"

-35_signal
-10_signal

gene
CDS
```


Yoshida,K., Yoshikawa,H.F., Zumstein,E., Yoshikawa,H. and Danchin,A.
The complete genome sequence of the gram-positive bacterium *Bacillus subtilis*
Nature 390 (6657), 249-256 (1997)
98044033
9384377
2 (bases 1 to 213680)
Kunst,F., Ogasawara,N., Yoshikawa,H. and Danchin,A.
Direct Submission
Submitted (18-NOV-1997) I. Moszer, A. Danchin, Institut Pasteur,
Regulation de l'Expression Genetique, 28 rue du Docteur Roux, 75724
Paris Cedex 15, FRANCE. E-mail: moszer@pasteur.fr,
adachin@pasteur.fr Phone: +33 (0)1 45 68 84 41, Fax: +33 (0)1 45
68 89 48

FEATURES
source
1. 213680
/organism="Bacillus subtilis"
/strain="168"
/db_xref="taxon:1423"
complement(90.782)
/gene="yocL"
complement(90.782)
/gene="yocL"
/function="unknown"
/codon_start=1
/transl_table=11
/protein_id="CAB13977.1"
/db_xref="GI:2634479"
/db_xref="SP:REMBL:O34581"
/translation="MDSKHKIKIINGYHKNICISCKWLPAENFYSVKKNKDGLHSY
CKACVLKAKESMLNYKQLEMRERNLLPGMKAKKYNSSLKKKKTQOIWOEKNK
LKNTLRQADAKKINIDVQWCKDYFNKKSCYGLKIEDHKILFKGTYYQSDFH
KEVHDHGANDISNCIPACKSCSKKHDFAFEWYNSNKNFSSERLLKIKEWLNRFK
EERQDSEWRTKG"
complement(822.1025)
/gene="yocK"
complement(822.1025)
/gene="yocJ"
/function="unknown"
/codon_start=1
/transl_table=11
/protein_id="CAB13978.1"
/db_xref="GI:2634480"
/db_xref="SP:REMBL:O34326"
/translation="MKLIERTVEERLYIYDDIEARKHFVSELKPGWMVDNCWVG
DLNQIKQFYRFRKELTDDVYFNR"
complement(1045.1560)
/gene="yocJ"
complement(1045.1560)
/gene="yocJ"
/function="unknown"
/note="similar to hypothetical proteins from B. subtilis"
/codon_start=1
/transl_table=11
/protein_id="CAB13979.1"
/db_xref="GI:2634481"
/db_xref="SP:REMBL:O34359"
/translation="MKNPTMLKLDKLLAVIELITKENKYRITGGALCTDQACWCH
VHLKKRPHKIKIATPFKEQDKVSWQKRWKRLNVADEIINVEELDKYKVSGE
KPGEFSPAKMKQKNEYMIDHSEIAVAIVYDGSKSGTRNCLNIAKTKYILGHQLWLRHPDF
NFEIDITYFVG"
complement(1771.1965)
/gene="yocL"
complement(1771.1965)
/gene="yocI"
/function="unknown"
/codon_start=1
/transl_table=11
/protein_id="CAB13980.1"
/db_xref="GI:2634482"
/db_xref="SP:REMBL:O34339"
/translation="MKSILLPEPKITVKGKPINDESNTKVIAIYDREPEPEAFW"

```

/transl_table=11
/protein_id="CABL3986.1"
/db_xref="GI:2634488"
/db_xref="SPTREMBL:O34834"
/translation="MEVNYKTNQREKIRNGPYDELDTOMHILIDLINFKIGLKTRH
CFCGHPYEQIOWFEDVNIKEDQILELAEIAGREWKGLQISFKWARFSPIMFNWS
LVLSKRPNPEDPKYRLRSVEFFESYAANK"
complement(4491..4829)
/gene="yocqB"
CDS
complement(4491..4829)
/gene="yocqB"
/function="unknown"
/codon_start=1
/transl_table=11
/protein_id="CABL3987.1"
/db_xref="GI:2634489"

Query Match      5.4%; Score 64.2; DB 1; Length 213680;
Best Local Similarity 54.4%; Pred. No. 2.1e-08;
Matches 129; Conservative 0; Mismatches 108; Indels 0; Gaps 0;

QY 335 GCCGGTCATGCAGACATCCCCACGGGTACAGCGTCCCTGTACATTCCTCTCAAT 394
    ||||| ||| ||||| ||||| ||| ||| ||||| ||||| ||| |||
Db 160924 GCCTGTTGCGCGCATACACGACATCCACCAAAAATCCCGTCCATGTTTAT 160983
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 395 GACATCAGGATCCCGCGCTCTCACATGGGATACCGCGCGGAGACTGACGCTTC 454
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 160984 AACCTCAGGATACCGCATGTTGTTCCATACAGGCACTCCGACGCCATCGCTTC 161043
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 455 AGCCAGTACCATACAAACGCTTCATTTCCGAGGAGTACACACCATCGCATCCG 514
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 161044 AAGCAGGACAAAGCCAAAGCTTCTTTTTCAGATAGCAGGCTTCAATCGTAA 161103
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 515 GTAGCCGTAACGTGGGAAAGGCGACCTGCCATTAACACATCTCCGCTCATTC 571
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 161104 ATAAAGATCTTCAACACGCTTGTATTTCCAAGCATTAAGACTTGGTCTCCAGCC 161160
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 20
AE003927/c 13334 bp DNA linear BCT 15-JUN-2001
LOCUS
DEFINITION
Xylella fastidiosa 9a5c, section 73 of 229 of the complete genome.
ACCESSION
AE003927.1 AE003849
VERSION
AE003927.1 GI:9105783
KEYWORDS
Xylella fastidiosa 9a5c.
SOURCE
Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
Xylella.
REFERENCE
1 (bases 1 to 13334)
Simpson,A.J., Reinach,F.C., Arruda,P., Abreu,F.A., Acencio,M.,
Alvarenga,R., Alves,L.M., Araya,J.E., Baia,G.S., Baptista,C.S.,
Barros,M.H., Bonaccorsi,E.D., Bordin,S., Bove,J.M., Briones,M.R.,
Bueno,M.R., Camargo,A.A., Camargo,L.E., Carraro,D.M., Carrier,H.,
Colauto,N.B., Colombo,C., Costa,F.F., Costa,M.C., Costa-Neto,C.M.,
Coutinho,L.L., Cristofani,M., Dias-Neto,E., Docena,C., El-Dorri,H.,
Facincani,A.P., Ferreira,A.J., Ferreira,V.C., Ferro,J.A.,
Fraga,J.S., Franca,S.C., Franco,M.C., Frohme,M., Furlan,L.R.,
Garnier,M., Goldman,G.H., Goldman,M.H., Gomes,S.L., Gruber,A.,
Ho,P.L., Hoelsel,J.D., Junqueira,M.L., Kemper,E.L., Kitajima,J.P.
and Marino,C.L.
The genome sequence of the plant pathogen Xylella fastidiosa. The
Xylella fastidiosa Consortium of the Organization for Nucleotide
Sequencing and Analysis
Nature 406 (6792), 151-157 (2000)
JOURNAL
MEDLINE
20365717
PUBMED
10910347
REFERENCE
2 (bases 1 to 13334)
Simpson,A.J.G., Reinach,F.C., Arruda,P., Abreu,F.A., Acencio,M.,
Alvarenga,R., Alves,L.M.C., Araya,J.E., Baia,G.S., Baptista,C.S.,
Barros,M.H., Bonaccorsi,E.D., Bordin,S., Bove,J.M., Briones,M.R.S.,
Bueno,M.R.P., Camargo,A.A., Camargo,L.E.A., Carraro,D.M.,
Carrier,H., Colauto,N.B., Colombo,C., Costa,F.F., Costa,M.C.R.,

```

```

Costa-Neto,C.M., Coutinho,L.L., Cristofani,M., Dias-Neto,E.,
Docena,C., El-Dorri,H., Facincani,A.P., Ferreira,A.J.S.,
Ferreira,V.C.A., Ferro,J.A., Fraga,J.S., Franca,S.C., Franco,M.C.,
Fromme,M., Furlan,L.R., Garnier,M., Goldman,G.H., Goldman,M.H.S.,
Gomes,S.L., Gruber,A., Ho,P.L., Hoelsel,J.D., Junqueira,M.L.,
Kemper,E.L., Kitajima,J.P., Krieger,J.E., Kuranae,E.E., Laigret,F.,
Lambais,M.R., Leite,L.C.C., Lemos,E.G.M., Lemos,M.V.F., Lopes,S.A.,
Lopes,C.R., Machado,J.A., Machado,M.A., Madeira,A.M.B.N.,
Madeira,H.M.F., Marino,C.L., Marques,M.V., Martins,E.A.L.,
Martins,E.M.F., Matsukuma,A.Y., Menck,C.F.M., Miracca,E.C.,
Miyaki,C.Y., Monteiro-Vitorello,C.B., Moon,D.H., Nagai,M.A.,
Nascimento,A.L.T.O., Netto,J.E.S., Nhani Jr.,A., Nobrega,F.G.,
Nunes,L.R., Oliveira,M.A., de Oliveira,M.C., de Oliveira,R.C.,
Palmieri,D.A., Paris,A., Peloto,B.R., Pereira,G.A.G., Pereira
Jr.,H.A., Pesquero,J.B., Quaggio,R.B., Roberto,P.G., Rodrigues,V.,
de M. Rosa,A.J., de Rosa Jr.,V.E., de Sa,R.G., Santelli,R.V.,
Sawasaki,H.E., da Silva,A.C.R., da Silva,F.R., da Silva,A.M., Silva
Jr.,W.A., da Silveira,J.F., Silvestri,M.L.Z., Siqueira,W.J., de
Souza,A.A., de Souza,A.P., Terenzi,M.F., Truffi,D., Tsai,S.M.,
Tsuhaiko,M.H., Vallada,H., Van Sluys,M.A., Verjovski-Almeida,S.,
Vettore,A.L., Zago,M.A., Zatz,M., Meidanis,J. and Setubal,J.C.
Direct Submission
Submitted (02-JUN-2000) Organization for Nucleotide Sequencing and
Analysis, Bioinformatics Lab - IC/Unicamp, C.P. 6176, Campinas, SP
13083-970, Brazil
FEATURES
source
1..13334
/organism="Xylella fastidiosa 9a5c"
/db_xref="taxon:160492"
/clone="9a5c"
96..541
/gene="XF0872"
96..541
/gene="XF0872"
/notes="similar to SP|P21364 (percent identity: 30 %/query
alignment coverage: 87.3 %/subject alignment coverage:
74.5 %); identified by sequence similarity; putative; ORF
located using Glimmer/RBSfinder"
/codon_start=1
/transl_table=11
/product="outer membrane protein"
/protein_id="AAF83682.1"
/db_xref="GI:9105784"
/translation="MVSCRRHPRAKWKENAAQNTVKEGQDVAPLTVTYVYINDNVG
FELWGITKLSTAKTDASGKIGNVTKPTPLASQYHFGQANVTFEPFGVYGSQAK
SKFTALGSTDRVNLGNFRGAIGTIGLDMVDSNWFARLDARTFRSNSKAALSSSL
GLDQNRRLDPWSVFGIGARF"
complement(747..1532)
/gene="XF0873"
complement(747..1532)
/gene="XF0873"
/notes="similar to GI|4156104 (percent identity: 43 %/query
alignment coverage: 102.7 %/subject alignment coverage:
98.9 %); identified by sequence similarity; putative; ORF
located using Glimmer/RBSfinder"
/codon_start=1
/transl_table=11
/product="outer membrane protein"
/protein_id="AAF83683.1"
/db_xref="GI:9105785"
/translation="MKHFLICVATLLFVCGSPDSNDRLSIAATVAPHAETLEFVKP
TLAKQGLDIRVFNVDVQVQAKQIDVNFQTEPYLOAYNRSRNTHLITVGVGH
IEFGVASYRTALASIPKGAETAIPDPSNRSRALILHQAOLIKLQPKMLATOR
DILNPLFKFRELDAMLPRVLQDIDLVLINTNYALSAGLKPNQDALIENKDSFV
NYLVSRDNKDPFRVOKLAKALSPVEVKFDIDKKYGSAILPAF"
complement(1626..2333)
/gene="XF0874"
complement(1626..2333)
/gene="XF0874"
/notes="similar to GI|6459112 (percent identity: 48 %/query
alignment coverage: 91.9 %/subject alignment coverage:
99.1 %); identified by sequence similarity; putative; ORF
located using Glimmer/RBSfinder"

```


Salzberg, S.L., Delcher, A., Utterback, T., Weidman, J., Khouri, H., Gill, J., Mikula, A. and Bishai, W.
 Direct Submission
 Submitted (25-APR-2001) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA
 Location/Qualifiers
 1. .14243
 /organism="Mycobacterium tuberculosis CDC1551"
 /strain="CDC1551"
 /db_xref="taxon:83331"
 /note="clinical strain"
 complement(102. .767)
 /gene="MT2233"
 complement(102. .767)
 /gene="MT2233"
 /note="similar to GB:AL123456; identified by sequence similarity; putative"
 /codon_start=1
 /transl_table=11
 /product="IS1558", transposase"
 /protein_id="AAK46518.1"
 /db_xref="GI:13881921"
 /translation="MRSKIPDLQALREDDHDMCRHLAHLQDLDMIGALDEQ IEQMHFPFCARRELIAISIPGIGVASATVISEIGADPAWFFSAEHLASWRCPCGNH ESARKRHGARTGNQHLQVLECAVAATVDGYLREYRQVRKFGGFRSPAANK AITAVAHKLIVIIWHVLATGRPYQDLGADYFTTRMDPKERRRLVAKLEAQGLGVTL E PAA"
 complement(1152. .2540)
 /gene="MT2234"
 complement(1152. .2540)
 /gene="MT2234"
 /note="similar to GP:2687633; identified by sequence similarity; putative"
 /codon_start=1
 /transl_table=11
 /product="phospho-2-dehydro-3-deoxyheptonate aldolase, putative"
 /protein_id="AAK46519.1"
 /db_xref="GI:13881922"
 /translation="MNVVDIPIDQLPSLPPLPTDLRLDAALAKPAAQWPPADR ALAMRTVLESPPVTPVSELVRLQDLQAOVAKGEAPFLQGGCAETFMNDTEPHIGN VRALLQNAVLTICASMPVVKVARIAGQYAKPRASDIDALGLSYBGDMINGAPDAA AREHDPRLVRAYANASAMNLVRLTSSGLASHLVDNWRFEVTSFACARYEALA TEIDGRFRSACGAVDNRLOTAEIYASHEALVLDYERAMRLSDGEGEPQLFDLSA HTVNGERTRIDGAHFAFAQVIANPVGVKGNMPELAVYERLDPNPKPCRLTL VSRMGNKVRDLLEPPIVEKQVATGHQVWQCDPMHGNTHESSTGFKTRHFDRIVDEYQ GFEFVRHAGLPHGCIHVEITGENVTCLGGAQDISETDLAGRYETACDPLRNTQOSL ELAFLVAEMLED"
 complement(2631. .3137)
 /gene="MT2234.1"
 complement(2631. .3137)
 /gene="MT2234.1"
 /note="identified by Glimmer2; putative"
 /codon_start=1
 /transl_table=11
 /product="conserved hypothetical protein"
 /protein_id="AAK46520.1"
 /db_xref="GI:13881923"
 /translation="MRVYFDTEFDGHTIELISIGVVAEDGREYVAVSTEFDPERAG SWVRTHVLKLPSPASOLWRSRQOIRLDLEFLRIDGTDSTELWAWGYADHVALCOL WGPWMLPPTVPTRELRQLWEDRGCPMPRPDRVDHALVDARDQLRRFLITSTD DAGRGAAR"
 complement(3147. .4085)
 /gene="MT2235"
 complement(3147. .4085)
 /gene="MT2235"
 /note="identified by Glimmer2; putative"
 /codon_start=1
 /transl_table=11
 /product="conserved hypothetical protein"
 /protein_id="AAK46521.1"
 /db_xref="GI:13881924"
 /translation="MVALYGACISQGGRSFLEVFHWLQHDIVDRGLPLCLLVAVFV

LTFLVTRSFVRFHRRRAADGRPARWQPNRNHIGSVHIIHHVAGVVLVNIISGLFLVTL
 SVDGPEFTTAAISIFGVGAALVLDDEYALILHLSDDWEEDGRTSVDAVFAAVAGL
 LIMGLHPLIFFLPGVROGANVVLQTTLIAGILVLTPLAVVLLKGVMTGLIGMFVVV
 LLVGCASRLSRPHAPWARMRYTRHPEKMRALQORERTWRPRVRKIKLVLQVVIAGTPR
 HPDERAVDAQLDQDVRPAPPERTAPILISGVMSD"
 4122. .5405
 /gene="MT2236"
 4122. .5405
 /gene="MT2236"
 /note="identified by Glimmer2; putative"
 /codon_start=1
 /transl_table=11
 /product="conserved hypothetical protein"
 /protein_id="AAK46522.1"
 /db_xref="GI:13881925"
 /translation="MSANRAPRSGRSLGRRVLMCLLWLAGVALGYVAVRLFHTPTVR
 IDIDIVQMCARAWLDGRPLYGCVLFTPTIGLNLPTYPPLAAVLFPFAMLOWPAAS
 VAITVLTLLIATAIYLTGLDAMPTSRYPAPARLRLVLAIVAPATWLEPIS
 SNFAGQINVLMTLIVDCPPRTWPRLMGLGIAKLTAPVFLYFLLRDGRA
 ALTALASAVATLLGLFVLAWRDSWEYTHLHHTDRIGAAALNTDQNALALRTIG
 DDERFALWVAGSLVLAATIWAMRRVLRAGETFLAVICVAFGLVVSFVSHHWMM
 LPAVLVIGLWRRNRNALMSLAGVLMRWTPIDLLPQHRETTAVVWRQLAGMSYV
 WWALAVIVAGLTVTARMTQBSLTRLGLTPAPTS"
 complement(5406. .6149)
 /gene="MT2237"
 complement(5406. .6149)
 /gene="MT2237"
 /note="identified by match to PFAM protein family HMM
 PF01553"
 /codon_start=1
 /transl_table=11
 /product="acyltransferase family protein"
 /protein_id="AAK46523.1"
 /db_xref="GI:13881926"
 /translation="MWYLFKVFIFMGPIFTLLGRPKVEGLEYPSSGPAILASNHVAV
 ADSEYLPVAVRRRIWFLAKSEYETGTGLKWLNRWFYSVSGOVPIIDRTNADSAQCALQ
 TAVLLGQCKLLGMPEGRSPDGLYKGTGLRLAHLTGVPVPIVAMIGTNVNNPP
 GRQMLRFGVTVFVKPDMFSRPEGLAGNHFIERAVTDEVIIELMGLSQEYVDIYAA
 SYKDGKRNAGAGANPNSTDAARIPETAAG"
 complement(6235. .6639)
 /gene="MT2238"
 complement(6235. .6639)
 /gene="MT2238"
 /note="identified by Glimmer2; putative"
 /codon_start=1
 /transl_table=11
 /product="conserved hypothetical protein"
 /protein_id="AAK46524.1"
 /db_xref="GI:13881927"
 /translation="MADVSGAHTDVRPELRKLAQAALDGDIPAVRVAAWASGGPGT
 GKQQVWCIPICALAALVTGEQHPHLTVIADHSLALLEVIRAIVDDIDRSKPPPEGPP
 GGGQTGAGGENTNGESMKSHYQAIPTVEE"
 complement(6627. .7769)
 /gene="MT2239"
 complement(6627. .7769)
 /gene="MT2239"
 /note="identified by Glimmer2; putative"
 /codon_start=1
 /transl_table=11
 /product="conserved hypothetical protein"
 /protein_id="AAK46525.1"
 /db_xref="GI:13881928"
 /translation="MLVSTDQAHSLGDVLGIATVPPTGGQDPVRVLAYPEAGGGFLD
 ALALDTLALLEGRLHVETLDRFPFSGSELSSIAPEELCALPGIQEVLGHVAGELAA
 ARWRDIRIVDCASTADALRLMTLPATFGLYVERAWPRHRLRSIGADGRSAVLAELE
 RIRASVERLSLTLDGALYSAHLVLTPERVAAEAVRTLSGLSALMGVRSVEELVNLLE
 VODENVEYSLDPHPAFHYTAERIGQRAVLDDDATIGDVALVLPVPHAGEPIGPKA
 LGGLDSARRRQGSAPPGLQPIVDLESGLSGLASYIRLRLALPQLDPTILTLGRADD
 LIIVSGMRRRRLASVLRCTVIDAHLRGELTVFRFNPENWPT"
 complement(7886. .8320)
 /gene="MT2240"
 complement(7886. .8320)
 /gene="MT2240"
 /note="identified by match to PFAM protein family HMM
 PF01553"

FEATURE	source	source	repeat_unit	repeat_unit	gene	CDS	misc_feature
TITLE	Barrell,B.G.	Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence					
JOURNAL	Nature 393 (6685), 537-544 (1998)						
MEDLINE	98295987						
PUBMED	9634230						
REFERENCE	2 (bases 1 to 23400)						
AUTHORS	Parkhill,J.						
TITLE	Direct Submission						
JOURNAL	Submitted (11-JUN-1998) Submitted on behalf of the Mycobacterium tuberculosis sequencing and mapping teams, Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA Unite de Genetique Moleculaire Bacterienne, Institut Pasteur, 28 rue du Docteur Roux, 75724 Paris Cedex 15, France E-mail: parkhill@sanger.ac.uk						
COMMENT	On Jun 20, 1998 this sequence version replaced gi:2911087.						
Notes:	Details of M. tuberculosis sequencing at the Sanger Centre are available on the World Wide Web. (URL, http://www.sanger.ac.uk/Projects/M_tuberculosis/) CDS have been renumbered from the original cosmid submissions but the old gene designations are in brackets after the new gene numbers. Gene prediction was based on a Hidden Markov Model of TB genes implemented in TBparse (Krogh) supplemented with visual inspection of positional base preference in codons, especially where there is an increase in the observed/expected third position G + C. CAUTION: In some cases we may not have predicted the correct initiation codon. Where possible we choose an initiation codon (atg, gtg, or ttg) which is preceded by an upstream ribosome binding site sequence (optimally 5-13bp before the initiation codon). If this cannot be identified we choose the most upstream initiation codon.						
FEATURES	Location/Qualifiers						
source	1. .23400 /organism="Mycobacterium tuberculosis H37Rv" /strain="H37Rv" /db_xref="taxon:83332" <1. .988 /organism="Mycobacterium tuberculosis H37Rv" /strain="H37Rv" /db_xref="taxon:83332" /clone="Y270" 22587. .>23400 /organism="Mycobacterium tuberculosis H37Rv" /strain="H37Rv" /db_xref="taxon:83332" /clone="Y190" complement(115. .1469) /note="IS6110 insertion element" complement(115. .142) /note="Inverted repeat at end of IS6110; GAGTCCTCCGACACACCGGGCGGTCA" complement(157. .1197) /gene="Rv2167c" complement(157. .>1197) /note="Rv2167c" /gene="Rv2167c", (MTCY270.01), len: 346. Function: IS6110 transposase. PASTA best: TPA9_MYCTU p19774 putative transposase for insertion sequence (identical)" /codon_start=1 /transl_table=11 /product="hypothetical protein Rv2167c" /protein_id="CAA17494.1" /db_xref="GI:3242294" /db_xref="SP:PREM1.033355" /translation="AEMALAGQRRIKAGERDFKRVGFLGRARPASTLITRFTIADHQGHREGDGLRWGVEICTQTELGVPIAPSYDHIINREPSRELRLDEGLKEHSRYVHAAANYGVGARKVWLTNREGIEVTEVLTKLSTGRKARTTIGRTTADTATPATASVQVQRGGPAPNRLWADLTYYSTWAGFAYVAVTGDVARRLTIGRWASTWATSWLDALEQALWTFROEGVLDKDLVHHTDRGSOYTSIRSERLAEAGIOPSGVAGVSSYDNLAEATNGLYKKELIKPGPKWRSIEDVELATWDFWNNHRLIYQYCGDVPVPELEAAYTAQRQRPAAG" 888. .22687 /note="fragment designated v021. Does not represent a						

```

physical clone"
complement(1092..1418)
/genes="Rv2168c"
CDS
/genes="Rv2168c"
/notes="Rv2168c", len: 108. IS6110
transposase. FASTA scores: sptri008155|008155 HYPOTHETICAL
12.0 KD PROTEIN (108 aa) opt: 697 z-score: 903.5 E(): 0;
100.0% identity in 108 aa overlap. Tbpase score is 0.928"
/codon_start=1
/transl_table=11
/product="hypothetical protein Rv2168c"
/protein_id="CAAL17472.1"
/db_xref="GI:2911088"
/db_xref="SWISS-PROT:Q50686"
/translation="MSGSSRRVPPELRERAVRMAEIRGHDSWAALSEVARLIGV
GCAETVRKVRQAVDAGARPCTTIESAELRLRDNALRRANILKTASAFAAE
LDRPAR"
complement(1141..1146)
/genes="Rv2168c"
/notes="DRAI:TTTAAA. Site linking fragments L-T"
complement(1442..1469)
/notes="Inverted repeat at 5'-end of IS6110"
complement(1563..1967)
/genes="Rv2169c"
complement(1563..1967)
/genes="Rv2169c"
/notes="Rv2169c", len: 134. Unknown. Tbpase
score is 0.934"
/codon_start=1
/transl_table=11
/product="hypothetical protein Rv2169c"
/protein_id="CAAL17473.1"
/db_xref="GI:2911089"
/db_xref="SPTREMBL:O53503"
/translation="MPLSDHEQRMLQIESALYAEDEPKFASVVRGGFRAPARRLQ
GAALFIGLMVLSVAFKEMTIGSPILSVFGVVMFGVYAITGPRLSGRMDRGG
SAAGSRQRTKGAGGSFTSRMEDFRFRFDE"
complement(1973..1979)
/notes="possible RBS, aaggagg, for Rv2169c"
2217..2223
/notes="possible RBS, aggagg, for Rv2170"
2233..2853
/genes="Rv2170"
2233..2853
/genes="Rv2170"
/notes="Rv2170", len: 206. Unknown. Tbpase
score is 0.905"
/codon_start=1
/transl_table=11
/product="hypothetical protein Rv2170"
/protein_id="CAAL17474.1"
/db_xref="GI:2911090"
/db_xref="SPTREMBL:O53504"
/translation="MAIFLIDLPDMERRLDALTYVDYDMRYPRGTETLRAPWLE
HIRRRQWAAVVAETAAQEAADTALPSAELSNAPMLGVAYGYPAGQWQQQ
VVLQSGRPRLAIRTSYFELTELHLPRAQGLGEALRLRLAGDEDNVLL
STPTNGEDNRNRLYRLGFTDIIRGYHFAGDPRAFAILGRTLPL"
2938..2942
/notes="possible RBS, aggagg, for Rv2171"
2949..3632
/genes="lppM"
2949..3632
/genes="lppM"
/notes="Rv2171", len: 227. Probable
lipoprotein contains putative signal peptide and
appropriately positioned PS00013 Prokaryotic membrane
lipoprotein lipid attachment site. Has hydrophobic stretch
at C-terminus and also contains PS00225 Crystallins beta
and gamma 'Greek key' motif signature. Unknown but similar
to . Tbpase score is 0.895"
/codon_start=1
/transl_table=11

```

```

/product="lppM"
/protein_id="CAAL17475.1"
/db_xref="GI:2911091"
/db_xref="SPTREMBL:O53505"
/translation="MATRRRGMLAIAMLLVPLATCCLRVRSITISPPDLVSGEI
IAAKPRNSKDTGFDALGDVDFPSQYAVSNYDSGYVGSQVFDLFAELPOLANMN
IAAGVNLURLRNGNIVILEGRADLTYSVDADVELTVAPPAAVTNGDRIEPEVV
QMKLPQGVYSTMSAQARYTDPNTRSFAGIWLGIATAAFAGVAVLAWIDRDRSPRL
TASGDPPTS"
2991..3023
/genes="lppM"
/notes="PS00013 Prokaryotic membrane lipoprotein lipid
attachment site"
3162..3209
/genes="lppM"
/notes="PS00225 Crystallins beta and gamma 'Greek key'
motif signature"
complement(3629..4534)
/genes="Rv2172c"
complement(3629..4534)
/genes="Rv2172c"
/notes="Rv2172c", len: 301. Unknown. Tbpase
score is 0.905"
/codon_start=1
/transl_table=11
/product="hypothetical protein Rv2172c"
/protein_id="CAAL17476.1"
/db_xref="GI:2911092"
/db_xref="SPTREMBL:O53506"
/translation="MTLNTIALELVPPNLEGGKERAIEDARKVQVYSNASGLGRIRH
VMPGMIADDDRRIPMPQPKLDVDFNSIIKPELAGVHGLCTQVAFMDEPSLHRRLLV
DLSDAGMEGIYFVGPRTMDGEGSVAPTDALSLYQLVANRGVIVIPTRDGGRL
NFKSRGATGYMTQLLYSDAIVGFLREFARTLHREPELLSFGVPKVEIRGLINWL
IODPCNAVADEQAFVOKLAGSEPARRRRLMVDLYKRVLDGVADLGPPLSLHLEATYG
VSAARFETFAEMLAYWSAPFKPD"
Query Match 4.8%; Score 57.2; DB 1; Length 23400;
Best Local Similarity 54.2%; Pred. No. 2.5e-06;
Matches 116; Conservative 0; Mismatches 98; Indels 0; Gaps 0;
QY 272 TTCTCTGGCGTGGCTGCATCCGGAAGCGTTCCGGTGGGATAAAAATCCGCAG 331
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 20072 TTCTCGCGCGCGCGCCATCGCACGCGCGCGATCCGATCGATCGATCGAC 20131
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 332 TCGCGCGTCCATGACAGACATCCCCACCGGTAACAGCGTCCCTGTACATTTCTG 391
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 20132 GCGTCCGCGCACCCCGTCCACCGACTACCGTCCGACCATAGCCCGCTTGTGCTG 20191
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 392 AATGACATCAGGATCCGCGCTCTCACTGCGGTAACGGCAGCGGAGACTGACGC 451
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 20192 CACCGTTTCCGCGTCCGCGAGAAATGCCGCGGATTACCGGCACGCGCGGAGGC 20251
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 452 TTCAGCGAGTACATACCAACGCTTCATTTTCC 485
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 20252 TTCGAGGACACGATGCCCAAGCCCTCGACGTC 20285
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
RESULT 23
AF498417/c
LOCUS
DEFINITION
Pseudomonas aeruginosa serotype O6 putative O-antigen biosynthesis
gene cluster, partial sequence.
ACCESSION
AF498417 AC104736
VERSION
AF498417.1 GI:20560066
KEYWORDS
SOURCE
Pseudomonas aeruginosa.
ORGANISM
Pseudomonas aeruginosa.
Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
Pseudomonas.
REFERENCE
1 (bases 1 to 14734)
AUTHORS
Raymond,C.K., Sims,E.H., Kas,A., Spencer,D.H., Kutayavin,T.V.,
Ivey,R.G., Zhou,Y., Kaul,R., Clendenning,J.B. and Olson,M.V.
Genetic variation at the O-antigen biosynthetic locus in

```



```
/db_xref="GI:20560077"
/translaton="MTTVKVIHVIAGLKVKVGAELMLKRLIETQNGSGBEFHSIISLS
DLGEGGLLEAGISVDTQSMRPMRVLRLIWIIFRERPDIVQTMWYHSDLLGG
LAARMGIGIIVNGFTTDLQEGGKSTVLVRKSGCAWLSGFLPKYIVCAEASRRRHI
AVGYNASRLVINGFTTDLQEGGKSTVLVRKSGCAWLSGFLPKYIVCAEASRRRHI
FVAAGLLARYSLKFLVGLRLLSSNAELQRLIETATGYAERFILLGRODVASCLK
AMDFCHSRTGEPGNVIGMAGWGLPCTITDVGDAAYLLGNDGVVVVPPALDNLGKG
IEDIALDVEGRGLPAGVAGVAGVAGVAGVAGVAGVAGVAGVAGVAGVAGVAGV
10568..11698
/notes="ORF 12; similar to Glycosyl transferases group 1"
/codon_start=1
/transl_table=11
/protein_id="AAM27823.1"
/db_xref="GI:20560078"
/translaton="MKPVLLIMVDPAFMHSRLPVAQAQAQGFVHATRPGDAVK
KLIVKQGLHVELSGKKNFSELYLLTVWRLMLRPDLVHLTVIKPVIYGGTAA
RLAPKGVGAAGSLGPFVMAKGLKACAFACVAVLYRRLGKLLKLVIFQNPDRDA
LIGLGTAFKSVLRSGVDLQYQAPESPETPVVTLAARLLRDKGVLEFVAAANI
LQKSPVAFOLVDLDDPGNPTSELESELRWSEGTIECLGYRDIASVFARSHIVV
LPSTREGLPKVLEASACGRAVTVTPVDCRDAIQADVTGLLVPRVDSAAALAIQRL
IESPELRKMGAAAGLAERDAERDAIESVQOHLDIYRAGLSGA"
11839..12657
/notes="ORF 13; similar to NAD dependent
epimerase/dehydratase family"
/codon_start=1
/transl_table=11
/protein_id="AAM27824.1"
/db_xref="GI:20560079"
/translaton="MLLGEISVWDAESADTVVHCARVHVMSETASDPLVERK
ANQVTLDLAREAVSRGVRRFIFISSIKVNGEGTEPGRTYADSPNPVDPYGVSKRE
AEQALLDABETGLEVVIIIRPVLYGVGVKAVOTMRLKRGVPLPLGAIHNRSLV
SLNDVLITTCIEHPAAVQVFLVSGDELTSTELLRRMGALGAPARLLVPASWI
GAARVLRQAFARLRGSLQVDIMKTROVLGTPPVGVQDALEKTAARSFLRQ"
12654..13694
/notes="ORF 14; similar to Glycosyl transferase group 4;
potential multiple membrane spanning domains."
/codon_start=1
/transl_table=11
/protein_id="AAM27825.1"
/db_xref="GI:20560080"

Query Match
Best Local Similarity 4.88; Score 57; DB 1; Length 14734;
Matches 144; Conservative 0; Mismatches 145; Indels 0; Gaps 0;

QY 234 TGTCCGCTCTGTTAATATCAACCGGTACTACTAATATCTCTGCGCTGGCTGCATC 293
Db 11655 TGGAGGATACCTTCGATCGCAAAATCAGCTCTGCCAACCTCGCCAGCAGCTCCATC 11596

QY 294 ATCCGGAAGCGTTCGCGTGGGATATAAAATCGCGAGTGCCTCCGCTCCATCGACACCA 353
Db 11595 TCTTGTGCAACTCTGCGGACTCAATAAGACGCTGAATGCGCTCCGCGAGCAGCAGAA 11536

QY 354 TCCCGCCAGCGGTACACGCTCCCTCTGCATCTTCTGATGATGATGATGATGATGATGATG 413
Db 11535 TCACGCGCGGACGACGAGCGCGCGTAAACATCAGCTGCGGATGATGATGATGATGATGATG 11476

QY 414 GTCTCACTGGCGATAACGCGGACGCGGAGTACGCTTACGCGAGTACATACCAACAAAC 473
Db 11475 ACATCGGTCTGCAACGCGGACGCGGCGCGGACGCGGCGCTGCGAGGAGCTTAGGAAGC 11416

QY 474 GTTTCATTTCCGAAGCGATGACACACACACTGCGCAATCCGGTAGACCG 522
Db 11415 CCTTCAGGATAAGAGGACGACGAGCTATATGCAAGCGGCGGAGACAG 11367
```

```
RESULT 24
AF035937/c
LOCUS AF035937 16736 bp DNA linear BCT 12-JAN-2000
DEFINITION Pseudomonas aeruginosa strain IATS O6 RpsA (rpsA) gene, partial
cde; Irf-Beta, Wzx (wzx), and Wxx (wxx) genes, complete cds; and
wbp gene cluster for O-antigen biosynthesis, complete sequence.
ACCESSION AF035937
```

```
AF035937.1 GI:6690121
Pseudomonas aeruginosa.
Pseudomonas aeruginosa
Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
Pseudomonas.
1 (bases 1 to 16736)
Belanger, M., Burrows, L.L. and Lam, J.S.
Functional analysis of genes responsible for the synthesis of the
B-band O antigen of Pseudomonas aeruginosa serotype O6
lipopolysaccharide
Microbiology 145 (Pt 12), 3505-3521 (1999)
20090474
10627048
2 (bases 1 to 16736)
Belanger, M. and Lam, J.S.
Direct Submission
Submitted (26-NOV-1997) Microbiology, University of Guelph, Guelph,
Ont N1G 2W1, Canada
Location/Qualifiers
1..16736
/organism="Pseudomonas aeruginosa"
/strain="IATS O6"
/serotype="O6"
/db_xref="taxon:287"
<1..805
/gene="rpsA"
<1..805
/gene="rpsA"
/codon_start=2
/transl_table=11
/product="RpsA"
/protein_id="AAF23994.1"
/db_xref="GI:6690127"
/translaton="XPKLEEGVEGLVHSEMDWTKNKHPSKYVQVGVGEVQVLDID
ERRRISIGIKOCKSNPWFESSQNKGRISGSISITDFGIFGLDGGIDGLVHLS
DISWNEVEEAVREFKKGDELEFVILSVDPERISIGIKQLEDPPFNSYSLHEKGS
IVRTVKEVDKAGVISLDDIDIEGLIKASEISRDRVEDARNLKEGEVEAKIISDR
KSRVLSLVKSKDVEDKDKAMKELRKQEVESAGPTTIGDLIRAQMENQG"
942..1226
/codon_start=1
/transl_table=11
/product="Inf-Beta"
/protein_id="AAF23995.1"
/db_xref="GI:6690128"
/translaton="MTKSELIERIVTHOGQLSAKDVELAIKTMLEOMSQALATGDRIE
IRGFGSLSLHPRAPVGRNPKTGESVRLDGKVFVHFKEKELRDVNEPE"
1954..2811
/gene="wzz"
1954..2811
/gene="wzz"
/codon_start=1
/transl_table=11
/product="Wzz"
/protein_id="AAF23996.1"
/db_xref="GI:6690129"
/translaton="MSGLWSRLWLLALLAVLPVLLAIVLFTTPVYEARVSTLPPSL
SDIAGFNLSRSENGLPASANGVYSIFVNLQSETRRQFEDVYIPSLSEASVGSR
DOLYALLMNSVTIQPKSEDSERYVIVIERHDPQSAEWGKEYIGIAVKRSMQEMQEG
VREFOVKNRNIIQQQIDILRETAKARREDRISRLKEALRVASAKLENPLITEGQVQDQ
QLSSIMSEGLMTMGAKAIRAEIKTLDQVSDDPFIPALRGLOEKYALLEGITLNLDK
VAVFRQDGR"
3064..4332
/gene="wbpO"
3064..4332
/gene="wbpO"
/codon_start=1
/transl_table=11
/product="WbpO"
/protein_id="AAF23997.1"
/db_xref="GI:6690130"
/translaton="MKDLKVAVVGGLYVGLPLAVEFGKRTVGVGFDINOGRIAEIQRG
IDSTLEVDAAELKEASELSFTNLQDKCNFVIVTPTIDEHKQDPLTPLVKASES
```

```
IGKVLKGDIVIVESTYVPGATEEDCVPLVEKFSGLRFNEDEFAGYSPERINPGDKEH
RVSSIKVTSGSTPEIAELVDSLIREIITAGTHKASSIKVAEAAKYVIENTQDLNIAL
INELALIFRNGDVEILKAAAGTKWNFMPPRGLVGGHCIGVDYPLTHKQASIGYH
PETILLAGRLNDMGAYVVSQVLKMKRRIHVDGARVLLGLMTFKENCPDLRNTKVY
DIVRELAENYNIOVDVDPWVSADAMHEYGITPVGTPSHGAYDGIILAVAHSEFKNMG
ASNIKRLGAERHVLVDLYLLDEKSDSLR"
4370. 5395
/genes="wbpP"
/genes="wbpP"
/genes="wbpP"
/codon_start=1
/transl_table=11
/product="WbpP"
/protein_id="AAF23998.1"
/db_xref="GI:6690131"
/translation="MMSRYEELRKELPAQPKVWLITGVAGSIGSNILETLKLDQKVV
GLDNFATQHRNLDVRSIASEKSNFKFIQODIRNLDCCNACAGVDYVLRHQALG
SVPRSINDPITSNATNIDGFLNMLIAARDAKQSFYTAASSTYDHPGLPKVPTIG
KPLSPVADKYVELYADVFSRCYCFSTIGLRYFNVFGRQDPNGAYAAVPIKWTSSM
IQGDDVYINGDETSDFCYIENTVOANLANTAGLDARQYVNIAGVGTSLNQLF
ALRDGLAENGVSVTHRPVYVDFREGDVRHSLADIKAALKLGYAPKYDVSAGVALAMP
WIMFLK"
5392. 5589
/genes="wbpQ"
5392. 5589
/genes="wbpQ"
/codon_start=1
/transl_table=11
/product="WbpQ"
/protein_id="AAF23999.1"
/db_xref="GI:6690132"
/translation="MIFTLVRLDESSVMSGAIFAFESRLVLQVLNIRGACDSLFSF
RSTFSMIVLFFTHCFLEFVW"
5596. 7020
/genes="wzx"
5596. 7020
/genes="wzx"
/codon_start=1
/transl_table=11
/product="Wzx"
/protein_id="AAF24000.1"
/db_xref="GI:6690133"
/translation="MLGKSLVYFLFKSPAILTLVGLSVFTRLLSPGYGYSLTII
VGLFNTVQLQWALGVGRYLPEDSQARARLLGPARAISLVLSVLIIFVTFLLMEW
REIGFSILYVGVGFLCAQWHDNLKIONALLOPLTYGKMLLIKAGSFFIGVLLV
YFGVGDVLLGLTFLSVLATIFFQDAWGRVSWALVDKQLTRLFAYGAPLITFLFA
FIVNASDRFFICAGFLGDAAGVYVSVDLAQXSVGTVAHVHLAFLPLVNEKLSKGL
POTQDLRFTFIFIAVVSFAAGLAWAPELSGSTMGEFREGALKIIPLLISLAFI
GALKSPYDYSQLASATRVQVYVAVAVDVFNLILIPFEGIVGAUVSSVMAFSS
ALISIFLGRVFPFALPGKDMKIALSVLLMAVSVASFSLAESAFFGLVVRVVLGGG
VYLAAMIALDVSGMRTFLKSLIR"
7142. 8188
/genes="wbpR"
7142. 8188
/genes="wbpR"
/codon_start=1
/transl_table=11
/product="WbpR"
/protein_id="AAF24001.1"
/db_xref="GI:6690134"
/translation="MTAELPAVNVRAFTQGRSVPSRFRVKALLEGIGRGMHMSSE
ALVSYAPPDARLDRPWLIRELLHRLPQVIGSDVILQIRELLSTIPTLEFLTKAP
RILDVDATLWLRGIAANSIARRVDHIVCGNOYLADYFGQFRPTTIIPTGVDTLRF
LPRRERNRIVGWSGTSGYRFLYDIEIPLSHLRDHPKMLRIVSDRPPKVP1PA
EKVEFTWETNEVETIAGMDIGIMPLADDLRSRKCYSKMLLYMACGLPVPVSEYGM
NRDVLARGIYGAVDDEGRYESLAALVNDPEARVRAGNGRDIIRHYSLDVVCOLW
AMVIGVAPGKGR"
8193. 10076
/genes="wbpS"
8193. 10076
/genes="wbpS"
/codon_start=1
/transl_table=11
```

```
/product="WbpS"
/protein_id="AAF24002.1"
/db_xref="GI:6690135"
/translation="MCGIAGFWNITGTLGDNARVAROMAAAIHHRGPDESIGIWEAP
RAPILVHARLAVLELSPAGSQPMHSDGVRVLYINGEYIYNHLARLRISEAGVTHSHR
GSDTEFLTACFAQMGVESTLKLTVGMFALALWDQEKITILARDMGKPLTWGQHN
GVLPFASSELKAKEHPLFRGDDIDRALALFLRGYYPAPISITKGLGKURAGSILVLS
ERSLNETCEPAAYSANAAIEEALSNPFGTDAEAVDLLESQRLTSISQOMSDVPLG
AFISGGVDSSTVALMQQSSRPITFISIGFDEPQDEAVYAKAAVAEHGICTDHTELYV
NSKDALDVIPLPKIYCEPFGDSSQIPLTIVGLARQOVTVASGCGDELPGYNPY
OFTPRVRLMERPHSMRPFASAFODLPLPEKLGKLRDVFASRTAEELFYRLNSHWR
NHEYVIGAGHTALDTPRWRPVDVDFQWMMMDVQGMEDDILYKVDRAAMNSL
ETRVPLIDHRYFELAWRMPLHMRKNGKWLLEVIYRHVSRELLERPKKFSVPVS
DWLGRPKWAESLDERLQOEGYLDLSRLIRIWNHLAGRDRHRSRLWSVLMFOAM
LES"
10073. 11224
/genes="wbpT"
10073. 11224
/genes="wbpT"
/codon_start=1
/transl_table=11
/product="WbpT"
/protein_id="AAF23993.1"
/db_xref="GI:6690126"
/translation="MTTVKVIHVIAGLKKVGGAEMLMKRLIETQNGSGSEFEHSIISL
DLGEFGGLIEAGISVDVIGMTSMRDMRVLRLIWIIFRRERPDIVOTMYHSDLLGG
LAARMAGIGGIWGVRTDLQEGKSTTVLRKVCWLSGFLPKYICAAEASRRSHI
AVYNAAGRLMVLIPNGPDLRLQATDQSGAISRGAGIEASDIVIGSLGRFHPVKDRAS
FVAAGCLLAPYSLRLKLLVGLRELLSSNAELORLEATGYAERFILLGRQDVASCLK
AMDFCLHSRTGFPNVLGEAMMAGLPCTTVDGDAAYLLGNDGVVVVPALDPNALGKG
IEDLIALDVGRRALGEAAQRISYNSFTMASQRSFMSLYRDILVLRKA"
11268. 12398
/genes="wbpU"
11268. 12398
/genes="wbpU"
/codon_start=1
/transl_table=11
/product="WbpU"
/protein_id="AAF23992.1"
/db_xref="GI:6690125"
/translation="MKPVLIMVNDPAFFMSHRLPVAVGAQQAQGVHATIRPGDAVK
KIVSQFLLHHELPLSRSGKNPFSLEYLLVYWRLLWLRDPDLVHLTIKRVYGGIAA
RIAPVAVVAASVGLGVFMAGLKACACRACVAVWLYRRALRKLRLRVIPQNPPDRDA
LGLGATTEKSVLIRGSDVLTQYQAPESPETPVYTLAARLLRKLGVLEFVEAANI
LQRSYPAHFQVLVDLPNGNTIESELEMRWSEGTIECLGRQDIASVFARSHIVV
LPSYRGLPKVLVEASACRAVYTTDVPQCRDAIQADVTGLLVPRVDSALADAIOBL
IESPELRKMGAAGRALAEERDPAIESIVQQHLDIYRALGSA"
12395. 13357
/genes="wbpV"
Query Match 4.8%; Score 57; DB 1; Length 16736;
Best Local Similarity 49.8%; Pred. No. 2.7e-06;
Matches 144; Conservative 0; Mismatches 145; Indels 0; Gaps 0;
QY 234 TGTGCGGTTCTGTTAATATCAAAACCGGTACTCAATATCTTCTGTGGCGTGGCGCATC 293
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 12355 TGGACGATACTTTTCGATCGCAAAATCACGCTCTGCCAAACCTCGCCCGCAGCGATC 12296
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 294 ATCCGGAAGCTTCCCGTGGGATAAAAATCCCGCAGTCGCGCGGTCCATGCAGACACA 353
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 12295 TTCTTGGCAACTCTTGGAGCTCAATAAGACGCTGAATGCATCCCGCAGACGACGAGAA 12236
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 354 TCCCCCAGCGGTAAACAGCGTCCCTGTACATCTTCTGAATACATCAGGATCCCGCC 413
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 12235 TCACGACCGGACGAGAGCGGCCGCTTAACATAGCTGATTCATCCCGGCAACCGGT 12176
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 414 GTCTCACTGGCGATAGCGGCACCGCGGAGACTGACGCTTTCAGCGTTACCATACCAAC 473
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 12175 ACATCGTCTGTCACACCGCACCGCCGCGACGCGCTCGACCGACCTTAGGAAGC 12116
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 474 GCTTCATTTTCCGAAGGATGACCACTGGCACTGGCAATCCGGTAGACCG 522
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 12115 CCTTCAGTAAAGAGCAGACGACTATATGCGAAGCGGCGAAGACAG 12067
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

RESULT 25
 AP003599/c
 LOCUS 237221 bp DNA linear BCT 28-NOV-2001
 DEFINITION Nostoc sp. PCC 7120 DNA, complete genome, section 19/19.
 ACCESSION AP003599 BA000019
 VERSION AP003599.1 GI:17134317
 KEYWORDS
 SOURCE Nostoc sp. PCC 7120 DNA.
 ORGANISM Nostoc sp. PCC 7120
 Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
 1
 Kaneko, T., Nakamura, Y., Wolk, C. P., Kuritz, T., Sasamoto, S.,
 Watanabe, A., Iriguchi, M., Ishikawa, A., Kawashima, K., Kimura, T.,
 Kishida, Y., Kohara, M., Matsumoto, M., Matsuno, A., Muraki, A.,
 Nakazaki, N., Shimpo, S., Sugimoto, M., Takazawa, M., Yamada, M.,
 Yasuda, M. and Tabata, S.
 Complete genomic sequence of the filamentous nitrogen-fixing
 cyanobacterium Anabaena sp. strain PCC 7120
 DNA Res. 8 (5), 205-213 (2001)
 21595285
 REFERENCE 2 (bases 1 to 237221)
 Kaneko, T.
 Direct Submission
 Submitted (02-MAY-2001) Takakazu Kaneko, Kazusa DNA Research
 Institute, The First Laboratory for Plant Gene Research; Yana
 1532-3, Kisarazu, Chiba 292-0812, Japan
 (E-mail: kaneko@kazusa.or.jp,
 URL: http://www.kazusa.or.jp/cyanobase/
 Tel: 81-438-52-3935 (ex. 2338), Fax: 81-438-52-3934)
 FEATURES
 source location/Qualifiers
 1. 237221
 /organism="Nostoc sp. PCC 7120"
 /db_xref="taxon:103690"
 /note="synonym: Anabaena sp. PCC7120"
 complement(72..734)
 /gene="all5179"
 complement(72..734)
 /gene="all5179"
 /note="ORF_ID: all5179"
 /codon_start=1
 /transl_table=11
 /product="O-methyltransferase"
 /protein_id="BAB76878.1"
 /db_xref="GI:17134318"
 /translation="MTTRIGITPNLYLLSVLSREPEILAQRLQETALQPMGRMOI
 APQGGFVALLVGLAKKTLEVGFTGYSLLVALALPAEGLKLVACDINEETALAQ
 RYQKAGVDHKIDLYLAPALETLDKLLVAGEAETDFEAFIDADKSNYDNYERSQLI
 RSGGVTAIDNVLWSGVADPEIQDNRTQIRAFNHKLQDQRTILSLPIGDLTLVR
 KN"
 867..1568
 /gene="all5180"
 867..1568
 /gene="all5180"
 /note="ORF_ID: all5180"
 unknown protein
 /codon_start=1
 /transl_table=11
 /protein_id="BAB76879.1"
 /db_xref="GI:17134319"
 /translation="MTFLGATDPRDPSYSDAARYLRIPAGTIRSWTVGRHYPIS
 NGSNFRPLIQICNLAPRLLSFTNVEHLVRAIRKHHQIDLGKVRDALDFIDEQFQI
 SHPLARFLTDGVDFIERYGLINASKSVKTELKDFNAHLREPDQDTGLAIKLY
 PFTFRSEENDPRFVVYDRIAFGRLVIGTGISTRVLAERYQAGESIDELAYDYDCDR
 LMLEEARCELCPTAA"
 1697..1996
 /gene="all5181"
 1697..1996
 /gene="all5181"
 /note="ORF_ID: all5181"
 unknown protein
 /codon_start=1

/transl_table=11
 /protein_id="BAB76880.1"
 /db_xref="GI:17134320"
 /translation="MEWPEVKGKGVVLTGDKISNNRLERIAVARAQIKWFTFASQ
 SILSGEMAGILLOAIVPMKRFVSKHPAPFIAKIYRDGHLDMKQAMLEELQEF"
 2453..3310
 /gene="all5182"
 2453..3310
 /gene="all5182"
 /note="ORF_ID: all5182"
 /codon_start=1
 /transl_table=11
 /product="oxidoreductase"
 /protein_id="BAB76881.1"
 /db_xref="GI:17134321"
 /translation="MAEKQTLPQQQQPPCTESKMQPKPOADDAOYEGSGKLDKVA
 LITGDSGIGRAVAIYAKAGADVAFVYLSEHDDAEETKNLVEQGRRAVSIGDITD
 EAFQRAIQCTVDFEGLDILINNAEQHPQESDITKEQLERTFTNIFSMFYLRK
 AAIHLKRGSAIINTSVYKYGSPOLLDYSTKGAIIVAFTRSLNLSIKGIRNAV
 APGIWTLPISTPFAEKVETFGQVPMQACQPEVAPSVVFLASDDSSYMSQVLLH
 PNGGEVNG"
 3552..3809
 /gene="all5183"
 3552..3809
 /gene="all5183"
 /note="ORF_ID: all5183"
 unknown protein
 /codon_start=1
 /transl_table=11
 /protein_id="BAB76882.1"
 /db_xref="GI:17134322"
 /translation="MSSLKPLQGTVELDCARANAKQGIETAAQYQCGYGDILNRPAGEL
 RKACEENLQVKELNELITDQDMILNLGTGTEIAPDTASEL"
 complement(3821..5221)
 /gene="all5184"
 complement(3821..5221)
 /gene="all5184"
 /note="ORF_ID: all5184"
 /codon_start=1
 /transl_table=11
 /product="Mg2+ transport protein"
 /protein_id="BAB76883.1"
 /db_xref="GI:17134323"
 /translation="MTETNNLNSTLDQVSRRELRLVTRQMRMLLEAGDLQAKAILV
 VPQADIAEAIETGPEMHALFSLKDEAIEVYEDYDLSVQERLEELKRSQVRDI
 VQSDSDRLARLDFELPAKVNNHLEQLSPTEQATALMLGYEADTAGRTMTLEFIAL
 KENMTIGALERIRSLANASMIYVLYTQDARLTGIVSLRELVTSPQOTIGEVT
 RDVFNVTTHQEEVARLIQRYFLAYVVDVQRLLVGIVYDVIDILEETFKDIY
 ALGGVQSGGNYFQMDLWEVARRWLFVLLITNVGTIIKSDQDILTKVNVTLTA
 FIPLLTGTGGNVGAQSSVTVIRGMNTDEIRSLGPLQVIGREATAGALLGMLGSIAT
 WAYFLQRIEVAIAVGASLIAISVLASISGSALEFLFRYLRLDPALMSAPFTTAVDV
 LGVLTIFNLARVILRL"
 complement(5361..6890)
 /gene="all5185"
 complement(5361..6890)
 /gene="all5185"
 /note="ORF_ID: all5185"
 /codon_start=1
 /transl_table=11
 /product="mercuric reductase"
 /protein_id="BAB76884.1"
 /db_xref="GI:17134324"
 /translation="MSNSELDRTVIRPVDVEYNOKLVAYVHPHPPQADNYDLVVI
 KAGLVNIPSOIDIDFPVAMRMVRVAGISHDNSAERFASLGVDFVLGSGRFASNTV
 EVACKTLKFKKVIATGARATPKAIGLEQAGYLTNETVPSLQRPKELAVIGGPGIG
 CELAQAFRRLGSEVYLHSGSHVNLKEDNDAAQIVQOTLKEGIRLVNLNKAEEVVT
 TEGRLYFSTNGHSDSVTDILVAGRSNVEGLNEAVGVYDKRRGVENDYDQT
 TNPKIYAAGSDICMDKFTADAAAIRIVIKNTLSPFGLGRSLSLVMPWVYTDPE
 VAHVGLYESQDQVETIKIPFSDVRAIADQDGFGLKHKKGSDEIVGATIVASHAGE
 MISEVTTAIVNKIGLNLSNVHPYPTQAEAKKAADYTRTLTLPTKLLGLTLTKF
 S"
 7074..7826

gene

QY	Db	Query Match	Best Local Similarity	Score 54.8; DB 1; Length 237221;
CDS		<p>7074. .7826</p> <p>/gene="alr5186"</p> <p>/note="ORF ID:alr5186"</p> <p>hypothetical protein"</p> <p>/codon_start=1</p> <p>/transl_table=11</p> <p>/protein_id="BAB76885.1"</p> <p>/db_xref="GI:17134325"</p> <p>/translation="MIPKQTRFRNRKRLHLLGLTALILIVARQFNICALLQTLILVQSLGFGPGPAYMILYLNATLLFIPGSLILTLKSGCLFGVFGWSVYVLIATVGAAILAFYGRILSRDWRVQIDIKYKPKFKMIDQSAKEGKWKVILTRLSPVFFPNLINTAFGVTCISUKDILSLGSLIPGTIMVYIGISLAGDLAKAGTNHQPVPTPETQIWQIMOGGLGLMATVGVYIITKIAKALSQKVVTEGIISKQDAE"</p> <p>7859. .8011</p> <p>/gene="asr5187"</p> <p>7859. .8011</p> <p>/gene="asr5187"</p> <p>/note="ORF ID:asr5187"</p> <p>unknown protein"</p> <p>/codon_start=1</p> <p>/transl_table=11</p> <p>/protein_id="BAB76886.1"</p> <p>/db_xref="GI:17134326"</p> <p>/translation="MLTTTDKIKHPFLGVNHEEKVFEQTLQLQLSLSLKPEKVASD ESHSHS"</p> <p>7989. .8657</p> <p>/gene="alr5188"</p> <p>7989. .8657</p> <p>/gene="alr5188"</p> <p>/note="ORF ID:alr5188"</p> <p>/codon_start=1</p> <p>/transl_table=11</p> <p>/product="two-component response regulator"</p> <p>/protein_id="BAB76887.1"</p> <p>/db_xref="GI:17134327"</p> <p>/translation="MNRILTAEDPRIAFIEKGLRSQGTTFATVATDAYSATNMALSSGFQDLIDGLPKGDLVDLEELRGGENIPVILTARDIDQKVAGFPAGADDYLTLPFPFELLVRVKARLQSGSOAMETVTKSGNIVLDRSRKVIQODTIELPAREFTLAPTFPHQGVLSRQLLDRWGYDYPGNSIVDVYVGYLRKKLGNLDLIEIVRGVGR LKPT"</p> <p>8884. .10497</p> <p>/gene="alr5189"</p> <p>8884. .10497</p> <p>/gene="alr5189"</p> <p>/note="ORF ID:alr5189"</p> <p>/codon_start=1</p> <p>/transl_table=11</p> <p>/product="two-component sensor histidine kinase"</p> <p>/protein_id="BAB76888.1"</p> <p>/db_xref="GI:17134328"</p> <p>/translation="MGQQAQKKEKSOVILFHPDPTOKSVKNOQNTIAGTFLKHRRGFPWS TPRILSWGIIILFFELFELFPIRQALYARNVORVNEEAIEKIEFEQLLGTNSIPK HQDESIEAIDKLTQDNRLKLPPTKQOLREFFDALGNQLPDDTFLIALMDKEFKYKSPARPKEMDRDAKLIRDAKLTQPKQEEVYIPNSQVDSIVYLARVEVQEQTGIM VVIATHTAGERDEVLETIAVIVQVSTFVLIFALVLAWSGKILAPLRLLTQTARKISETDLNORTIDNGBLAELGTTFNDMDRLQNAFIQSQNFINDAGHELRTPTIIRGLHLEMDPKIEIQETLTIVIDELEMRNRFVNDMLLAKAERPDLFLLLETFDVCFTTEE</p>	<p>275 TCTGGCGCTGGCTGCCATCATCCGGAGCGTTCGGCTGGCGGATAAAAAATCGCGCACTGC 334</p> <p>74530 TCGGGCGGCTCTCCCATGTCTGCTAATCTTTGGATGTGAATACACCATTCAGTTC 74477</p> <p>335 GCCGGTCCATGCAGACATCCCCACGGGTAACAGCGTCCCGTCCACATCTTCCTGAAT 394</p> <p>74470 TTCTTTGAGTGCCTTACATCTCGATGGGAATATGAACCCATCTACTCCATCATTAAT 74411</p> <p>395 GACATCAGGATCCCGCCGCTCTACTGGCGGATAACGGGACGCGCGGAGACTGACGCTTC 454</p>	

Db 74410 AATATCAGGCCCTGCTGTATTATATAAGTTGTAAATTACAGGAATTCACAGACATGGCTTC 74351

QY 455 AGCAGTACCATACCAAACGCTTCATTTCCGAAGCATGACCAACACACTGGCAATCGG 514

Db 74350 TAATAGAACATGCTCCAAACCCCTTCATTTAAGAAGGAATACTATAAACGCTAGCACTGCC 74291

QY 515 GTACACGGTAAACGCTGGGAAAAGGCACACTGCC 548

Db 74290 GTAATACTGTTCAGAAGTAATAATGGGAACTGAC 74257

RESULT 26
AE008327/c
LOCUS
DEFINITION Agrobacterium tumefaciens str. C58 linear BCT 18-DEC-2001
of 187 of the complete sequence.
ACCESSION AE008327 AE007870
VERSION AE008327.1 GI:15159773
KEYWORDS
SOURCE
ORGANISM Agrobacterium tumefaciens str. C58 (Cereon).
Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
Rhizobiaceae; Rhizobium.
REFERENCE
AUTHORS Hinkle,G., Slater,S.C. and Goodner,B.
TITLE Complete Genome Sequence of Agrobacterium tumefaciens C58
(Rhizobium radiobacter C58), the Causative Agent of Crown Gall
Disease in Plants
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 10197)
AUTHORS Hinkle,G., Slater,S.C. and Goodner,B.
TITLE Direct Submission
JOURNAL Submitted (14-AUG-2001) Bioinformatics, Cereon Genomics, 45 Sidney
Street, Cambridge, MA 02139, USA
COMMENT Approximately 800 bp of telomeric sequence missing from the left
end of the chromosome and 200 bp missing from the right end.
FEATURES
source
1..10197
/organism="Agrobacterium tumefaciens str. C58 (Cereon)"
/strain="C58"
/db_xref="taxon:181661"
219..1367
/gene="AGR_L_2541"
219..1367
/gene="AGR_L_2541"
/note="(117900) putative hexose transferase"
/codon_start=1
/transl_table=11
/product="AGR_L_2541p"
/protein_id="AAK89838.1"
/db_xref="GI:15159774"
/translation="WTSALFVSHTEKGAELEFLADLVKAGPHSRACFLSGCATAE
DLAPAGPVMLSAGREKMLSIIRNSFSGALARGAADVMVAWQLSREAKHFDVICA
NALFVCAAKLSRRPLWILHDITVDTFATSNRRASLAFARLAVLNSEETGR
AFTAEAGADKVRIVYNGDPAKALHDAGMAARLRAELGLGPQLVGLFGRLSEWKG
HYFTLDALAKMGVQAVIVGGALFEGOEAEARIQOASRLGLDGRVRLFGSPDVEL
MASDWMVAHTSIVAEFGFVVEAMCMGPRPVATRGGGVTEIIRGEGGLVYPGDAS
ALAAAGLTSLDPAQLRGQSGREDVSDRFSLQETCRSVSALLTEAA"
1396..2514
/gene="AGR_L_2543"
1396..2514
/gene="AGR_L_2543"
/note="mannosyltransferase B - Synechocystis sp. (strain
PCC 6803)"
/codon_start=1
/transl_table=11
/product="AGR_L_2543p"
/protein_id="AAK89839.1"
/db_xref="GI:15159775"
/translation="MPAGGRARNRILNMPQSYGKGASGVARTFSLERLELTGTD
YILRSPTWREQDPSRLEMETPRRIMVDFVLRQAAMPLICRRHGGIDVLFN
YDPFGSLPGKRRLLVHDLYFKTIPQIGWRATLTDFLRLMAGSNRVVCSEATR
KOLAREVPAATKSLIIFSDTSLSDPDTPVDATSTAGRYILAVGNATSNKWFALICR


```

CDS
<1..775
/gene="gmbd"
/codon_start=2
/transl_table=11
/product="ADP-L-glycero-D-manno-heptose epimerase"
/protein_id="AAL23753.1"
/db_xref="GI:16445332"
/translation="ENFLQIMAGEFEIEAIFHEGACSSSTTEWGGKYMMDNNYOYS
KEUHLCLEREIEPLFVASSAANTYGGRTSFEIESYEQPLNVYGISKFLFDEYVRIL
PEANSIQVFRFNVYGPREGHGKSMASVAFHNTQLNNGESPKLFGSDGFKRDFVY
VGDVADVNWENGWGSIFNLGTGRAESFQAVADATLAYHKRGSIETIYIPPEKLAGR
YQAFQADLTNLKAGYDKPKTKVAEGVTEYMAWLNRA"
785..1831
/gene="waaF"
785..1831
/gene="waaF"
/note="complements a Serratia enterica serovar Typhimurium
waaF mutant"
/codon_start=1
/transl_table=11
/product="heptosyltransferase II WaaF"
/protein_id="AAL23754.1"
/db_xref="GI:16445333"
/translation="MKILVVGPSWVGDMMSORLYRTLRAVYQAIIDVMAWCRPL
LQPMVNOALMPLGCHGALGIGERRLGRALRANRYDRAYVLPNSFKSALVFFPADL
PORTGREGNYGLINDVRVLDKAAFPMLVQRYVALAYDKGRVORADDLPQPLNLPOL
RYSDEIATTSAFNLTDSRPVYFCGAEFGPAKRWPHYHYAALAAQLPTESYQIVL
FGSAKDEAGEQIRAAALQDDARDFCLNLAGKTQLEQAVILIAACRAVVSNDGLMHA
AALNKALIALYGPSPDFTPPLSDKARVIRLSGYHKVRKGDAEQGYHQSILDIQPOQ
VLDALPLLVASEE"
1834..2799
/gene="waaC"
1834..2799
/gene="waaC"
/note="complements a Serratia enterica serovar Typhimurium
waaC mutant"
/codon_start=1
/transl_table=11
/product="heptosyltransferase I WaaC"
/protein_id="AAL23755.1"
/db_xref="GI:16445334"
/translation="MOVLIVKTSMDGVHLHTLPALDQAIPDIRFDWVVEGFSQI
PTWPAVDVPIVAIRWRKNWFSQDROCRDFKALQERVDVVIDAQLMKSAAAL
ITRVAGSKHGPKCKSARPPFASWFYNVRHEIDKQHAVERTELFAKSLGYKPKGSY
GDVAIAARLQSPVQVAGQVLFHATTRDDKHWPQONNREILATLADSLKIKLUPWG
AEHEQRALRLAEGFPFHEVLPKLSQQVADYLAGAVVSDVTGLSHLTAAALDKPNI
TLFGDTPGLIGYGQNHSLISPEKSMATIDATTAQALQKVIA"
2796..3962
/note="orf4"
/codon_start=1
/transl_table=11
/product="putative glycosyltransferase"
/protein_id="AAL23756.1"
/db_xref="GI:16445335"
/translation="MKLHIINLKGWGVVERFLQVINDTDCSNQVLCISGDIGBEI
RQGLPHQPTFANRLINALPKCPQFLKFLKWKIERANADVIVVDVLVGLAARP
RKGLVYDHGCSWRYPKNOKTRFFAMLDGVISASHASKRMELRFLPCPNHVVIN
RKTQAGTTLTQPIRIGTASRLVSLKISVSLAMQELLRRGHVDVLEVAGKG
PDRAFEALAAQLGDRVTFCGYQDDVAGFNRTHYNTPTITEPFGLSMESLYFG
VPVIFQVQDQPAVKGVCIGLVPSTVTEOHRQLTDIEVDFPHEVYDPLTDSLVSP
KLASHIDCADAVETLVAGETYQLLSRNAQRYPAEHFYAQFKAEPDDTLRSFIV"
3991..5232
/gene="waaL"
3991..5232
/gene="waaL"
/note="similar to Escherichia coli K-12 WaaL"
/codon_start=1
/transl_table=11
/product="lipid A core: surface polymer ligase WaaL"
/protein_id="AAL23757.1"
/db_xref="GI:16445336"
/translation="MLKOKTPHPAFSYLIYLGCAIAFCTIPFGSATGRNLFFVSSVIA
FIACVLPYRFNSFNALLLPALMFCVGMGTILMHHFKQPGGYIYIYSYMTGKLQ

CDS
/note="orf6; similar to TsrB (Wbck) protein from Versinia
enterocolitica"
/transl_table=11
/product="unknown"
/protein_id="AAL23758.1"
/db_xref="GI:16445337"
/translation="MKTHQPTLLSVIPFYNNFAVIAALDSLTQISDEIEVVIID
DDYDASGALVSRYLAHQHPRVAFISOANGIAHARNVGLQHAAGRYTLFDGDGLLS
GDTLAILRPLLIADVDLIDFNKTEQPKPENRATOPAAVDFSLGSLQPLFA
RSMWHLMSRYVKTLLAGETFEDEGRYEDVIFPFQYFKRRIAHLDCELYFTNDSQ
GITRNKPKDIEDLLFAMEKMLRFVAQHPQDEPLRLAALMLANCSEVKMSKAVYG
YHYAPATIQILRRATVCRNSDVPCCKVKQMYARVDYDTLSKLGRKPR"
5229..6209
/note="orf7; LPS from a non polar orf7 mutant has reduced
levels of D-D-heptose"
/codon_start=1
/transl_table=11
/product="heptosyltransferase"
/protein_id="AAL23759.1"
/db_xref="GI:16445338"
/translation="MSKKWKIKVADMFLERYTAKYGFQDKATFDCAAQOLDNVIYST
TALGDPMFNTPALRAMERYPNAHILVHVKKRELVENGRYDRVYVDSKNTIRP
LKLSEHQDPLALILHSHLPYDIISAVMAGAKYIINDNYSSGIGLEPWLTNYIFHY
YGHFVRKKLELISILGCRSDPTMEIPVDYQGVKSDHPTFQIGLAGSTECRWVYS
HYARLADDELVAKHPNVRIVLGLADPDAQQFMSLTGPQACVNDQIKGTQGLLL
TLDSFHLVLTGDTPLHLAVALKVKITISLFTVADSRTGPIQDPELLRIQVSRGTY
TMLEAESPMGVIOPIQRIYDELVAHVPLAR"
8516..9598
/gene="waaQ"
8516..9598
/gene="waaQ"
/note="complements a non polar Klebsiella pneumoniae waaQ
mutant; similar to WaaQ proteins from the
Enterobacteriaceae"
/codon_start=1
/transl_table=11
/product="putative heptosyltransferase III WaaQ"
/protein_id="AAL23760.1"
/db_xref="GI:16445339"
/translation="MMDAPALTSAQVQRILLIKLHHGDMLLVTPVVISLQNYPO
ARIDVLYQTEMLASNPESLTIFAIDRWKKQGAHGLHGLALLRLKRAQRYDILV
VNLADQWRSALITRLTGARIRLIGDFPKRRGFLWRHCHTQLVPVNNHAHLTHVQNL
LLAPNLPALNHEVTSYDPQDQTCQLLQKQVAGSYIVVQTSNWFKWCSEKRM
AATLTALQADGHOLVITSGPDAREKAMVERILALCPQGVISLAGOLTLRLAALIDH
AKLFIGVDSVPMHMAALQPCVALFGPSKLTFRWPQATGAVIAGDFGELPDPDAI
DTGTDERYLDLIPTDVIAAARSTLA"
9595..10722
/note="similar to Klebsiella pneumoniae C3 putative
glycosyltransferase and WaaG from Salmonella enterica
serovar typhimurium, and Escherichia coli strains K-12,
R2, R3 and R4; Orf9; LPS from double non-polar orf9-orf10
mutant lacks GalA, has reduced Glc levels and contains
O-antigen"
/codon_start=1
/transl_table=11
/product="putative glycosyltransferase"
/protein_id="AAD28801.1"
/db_xref="GI:4753135"
/translation="MKAFLLAIYRRKYRYPDGAERFVSRAKALEQQDLDLNVITREW
QGANPNWHIHLNPLKRLGRISRGFAVARALWQKREFDLVQSHREIPCCDIYRAG
DGVHREWLQRLALPEWRKKWLFNSRYHYRVMAERAMTAAPELXAVICNAEWIERE
IADFGVPADKITVIYNAIDNQKFPDAEQRRRLREYOYIQQOACHLCITFVSGFERK
GLAAATRAVAATDHLVVGDKDAEKYRALAOSLGGDRIFHEMGVOKOTLPFYQAD
ALLPLIYDFPNVILEAMSCGLPVITSTTTCGGAETITPCQGEFVTDALDVPALTEAI
RALPQALGSSMGAAKLRIWTATPAHLSEQLISLYNRLID"

```

Query Match 4.3%; Score 50.6; DB 1; Length 20693;
 Best Local Similarity 50.2%; Pred. No. 0.00034;
 Matches 125; Conservative 0; Mismatches 124; Indels 0; Gaps 0;

QY 382 CATCTCTCTCAATGACATCAGGATCCGCCCGCTCTCATCTGGCGATGAACGGGACCGCGG 441
 DB 11667 CTTTTCAGAAATTCGCGGAGCGCGCGGAGCGGCTGTCGACCAACCGCGTCCCGC 11608

QY 442 AGACTGACGCTTCAGCCAGTACCATACCAACGCTTCATTTCCGAGGCGATGACCAACA 501
 DB 11607 ACAACAGGATTCACCGACGATGCGCAACCCCTTCGCTATCGGAGCTGAGCAACCAACA 11548

QY 502 CACTGGCAATCCGCTAGACCGCTTAACGCTGGGAAAGGCGACCTGCCATTAACACATCTC 561
 DB 11547 GCGAGGATGCGGATTAACGGATAGGATGGCTGGAAACCGAGGACGACGCGCT 11488

QY 562 CGCTCATCTCCAGGTGTTCTGTCTGTCGACGACAGCTGCTTCTGTTATTTTCACGCGCGG 621
 DB 11487 CTGCGATACCAAGTCTGCCGCCAGACGCTTGACCTCGCGGAGCGGCGTCAATCCCGG 11428

QY 622 CGCCACCA 630
 DB 11427 TGCCGATCA 11419

RESULT 29
 LOCUS AE013034 8594 bp DNA linear BCT 09-MAY-2002
 DEFINITION Thermoanaerobacter tengcongensis strain MB4T, section 61 of 244 of the complete genome.
 ACCESSION AE013034 AE008691
 VERSION AE013034.1 GI:20515643
 KEYWORDS
 SOURCE Thermoanaerobacter tengcongensis.
 ORGANISM Thermoanaerobacter tengcongensis.
 Bacteria; Firmicutes; Clostridia; Thermoanaerobacteriales; Thermoanaerobacteriaceae; Thermoanaerobacter.
 1 (bases 1 to 8594)
 Bao, Q., Tian, Y., Li, W., Xu, Z., Xuan, Z., Hu, S., Dong, W., Yang, J., Chen, X., Xue, Y., Xu, Y., Lai, X., Huang, L., Dong, X., Ma, Y., Ling, L., Tan, H., Chen, R., Wang, J., Yu, J., and Yang, H.
 A Complete Sequence of the T. tengcongensis Genome
 Genome Res. 12 (5), 689-700 (2002)
 2 (bases 1 to 8594)
 Bao, Q., Xu, Z., Hu, S., Dong, W., Chen, Y., Wang, J., Yu, J. and Yang, H.
 Direct Submission
 Submitted (07-OCT-2001) Beijing Genomics Institute/Genomics and Bioinformatics Center, Institute of Genetics and Development, Chinese Academy of Sciences, Beijing Airport Industrial Zone B6, Beijing 101300, China
 3 (bases 1 to 8594)
 Li, W., Xuan, Z., Yang, J., Ling, L. and Chen, R.
 Direct Submission
 Submitted (07-OCT-2001) Institute of Biophysics, Chinese Academy of Sciences, Beijing 100101, China
 4 (bases 1 to 8594)
 Tian, Y., Xue, Y., Xu, Y., Lai, X., Huang, L., Dong, X., Ma, Y. and Tan, H.
 Direct Submission
 Submitted (07-OCT-2001) Institute of Microbiology, Chinese Academy of Sciences, Beijing 100080, China
 FEATURES
 source
 1. 8594
 /organism="Thermoanaerobacter tengcongensis"
 /strain="MB4T"
 /db_xref="taxon:119072"
 102..1289
 /gene="RfaG"
 /note="TTE0655"
 102..106
 /gene="RfaG"

117..1289
 /gene="RfaG"
 /note="Best Blastp hit = gi|2129096|pir||F64500 probable hexosyltransferase (EC 2.4.1.-) MJ1607 - Methanococcus jannaschii gi|1500505|gb|AA899629.1|'(U67601) LPS biosynthesis protein, putative [Methanococcus jannaschii]', score 100, E-value 3.00E-20"
 /codon_start=1
 /transl_table=11
 /product="predicted glycosyltransferases"
 /protein_id="AA023922.1"
 /db_xref="GI:20515644"
 /translation="MRDNRIQDVALLISNYPITNLGQLHLLLEKRLNNVNVYTKYTFSSFEKAKLVFPYFKMSYKDRVITRTKLEKFKQHFQVFNADVI SGACNSKIVLTHGYTRAINYGFSGEDIPDFERGLDEKRAIDQVAFIPR SRKGVVSEFSEPKEDVIFNAVDIDFKPISLEKKQIRELWGCEDFVVFIPR RYKNGVIVAAKAAIILKEDIKFIFAGIGLKEIILNTHKNNKVLKSGIDPEI VRYKACDVLIPSTSDGVEATSLSMLEMSCGKIYVCTPIGCKMELIKHGVNGFF VEQKSEAIAYIIIEKIDFYKLDISIRQEARKIIEKNHSYIVHARKFIEVYKAIR"
 705..1217
 /gene="RfaG"
 /note="Pfam match to entry glycosyl transferases group 1, score 123.5, E-value 3.90E-33"
 1350..2742
 /gene="TTE0656"
 1350..1354
 /gene="TTE0656"
 1369..2742
 /gene="TTE0656"
 /codon_start=1
 /transl_table=11
 /product="hypothetical protein"
 /protein_id="AA023923.1"
 /db_xref="GI:20515645"
 /translation="MFTLVILGLMIVFEIWRKIKFDLFLRGNIYVFIYEL ARSYIILDFKVDSLPPSSMIFENYKDAYNSFLASVISVIGFCTVGYLGLSIF RNIGCSINFKLNKISDSIWLIIIVLFGFSGLYIKRSAPFDIIMVGLLSRG IKVGEETSTFTSLVTFSSYLLGLSKNFSFKKFLYFLGFGISGLSLV LILKAGRLHNLVLLVFTLYSTLEKIRFLVLSLGSLSLILYKGYWFGIRFG KPIYNSLODIFSNFAEFSPYLSINILCNPKYISRYFLPDLPLRILPLGLISK IIDINPLTLQENTLIIILGTEKGPIDILGAGYFSANIFGVVTFMFCGLMLAL PERIFSNYSNPVLSRLKAWILYFSTMVLYADPAILLMDGVYLIFPIVLEMLYLL NEKKYSP"
 2751..3904
 /gene="RfaG"
 /note="TTE0658"
 2751..2755
 /gene="RfaG"
 2762..3904
 /gene="RfaG"
 /note="Best Blastp hit = gi|7433884|pir||E72354 probable hexosyltransferase (EC 2.4.1.-) TM0622 - Thermotoga maritima (strain MSB8) gi|4981142|gb|AAD35706.1|AE001736.4 '(AE001736) lipopolysaccharide biosynthesis protein, putative [Thermotoga maritima]', score 263, E-value 4.00E-69"
 /codon_start=1
 /transl_table=11
 /product="predicted glycosyltransferases"
 /protein_id="AA023924.1"
 /db_xref="GI:20515646"
 /translation="MNRKIKQIIPNFGYGAERTVNLKYLDEKKEVEYRAISMFD LNTLEKLENIENIPYVYLGKKGDPFRMFRIDIKISFKPHIVHTHYVLRVALPS LLLHKVPKVHVHNAEKEDVKGLVKAFISFGVTPISIRLVSESLISVYGVKN IPLILNGIPVEYQKANTNREWEKGEFQKEDFLFNARLAPKQNALIEAFAGK PARHNSKLIIVGDCGERLEITKHLREKUYELGIRWDIPDLNASDVVFLSSD WEGNLSKNEAAKAPKPIATSVGGVPELIQNNINGILVPPKNNVAFKAMLIENK DICQKLGEKAKEAEKEDFISVWKKYKELYESLLQFLKRGKASLL"
 3296..3802
 /gene="RfaG"
 /note="Pfam match to entry glycosyl transferases group 1, score 199.9, E-value 4.00E-56"
 3886..4839

```
/gene="TTE0659"
3886. .3890
/genes="Rfag6"
/genes="TTE0659"
3901. .4839
/genes="TTE0659"
/notes="Best Blastp hit = gi|13442978|gb|AAK26246.1|
(AF330049) D-glucuronyl C5-epimerase [Mus musculus], score
96.3, E-value 5.00E-19"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAM23925.1"
/db_xref="GI:20515647"
/translation="MYVRLSKYLILLGKSLAHVQNGIGKIYSKEGIVYNDLTL
KYNHGETKRGLPNMVMDGKIYPTTVIQGLGNIYDILITKQDITFEQWKSVIN
WVNEDEGLGMDVFERIGSKFKPKYSAMTQEGASLLIRAFQESGDEFVDRAKAI
DMLLPEVGGARYYNDKLFEEVYESKPLVNLNGWIFAVFLGYDIFRATSDGRKE
ALQRTLDLTKDELYKYDTGYWSYDQCGNLASPYHKLHIALLELVLYELFNTIEFTV
KQWESYLSNKKRTRAFIVKAYOKIKEPGEVLIK"
4840. .6072
/genes="Rfag6"
/notes="TTE0660"
4840. .4844
/genes="Rfag6"
4852. .6072
/genes="Rfag6"
/notes="Best Blastp hit = gi|7470136|pir|I577339
hypothetical protein sll1723 - Synechocystis sp. (strain
PCC 6803) gi|1652521|dbj|BAAL7442.1| (D90906) hypothetical
protein [Synechocystis sp.], score 117, E-value 3.00E-25"
/codon_start=1
/transl_table=11
/product="predicted glycosyltransferases"
/protein_id="AAM23926.1"
/db_xref="GI:20515648"
/translation="MKGMIAYITSTPYKGGQFQILPELLEVIKKGHDVVPVPE
KETALDEPKVEAKFMHPIPLLSLILKSLFIFKHLRVVVEILFKILRHSGSLGKI
LKNLLIPIKGLVSEILKKRKNIHIAHWASPTLAYIASELTGIPWFAHRWDIA
ENMLREKARSNFIRVDKNGYDELSVKEYTRDKLYILHVGKRVYVNFKLDLAKKA
NRALLAVANFVYKGGHLYIQAIKKVITQNCVYCYFFGKGLLEDLKLTVTQLN
LEDRIFFQGLPHELDIYKIDCVILPSIVTDGKEGIPVSLMEAMAHKIPVI
STNGGIPELLEGGAGIIVEQKNSDELAKINLINDEXLELGEKGEKIEKFNEL
SKIVEBLLVMGN"
5470. .6006
/genes="Rfag6"
/notes="Pfam match to entry Glycos_transf_1, Glycosyl
transferases group 1, score 154.5, E-value 1.80E-42"
6085. .7190
/genes="WecB2"
/notes="TTE0661"
6085. .6089
/genes="WecB2"
6099. .7190
/genes="WecB2"
/notes="Best Blastp hit = gi|3451512|emb|CAA07668.1|
(AJ007747) putative UDP-N-acetylglucosamine 2-epimerase
[Bordetella bronchiseptica], score 381, E-value 1.00E-105"
/codon_start=1
/transl_table=11
/product="UDP-N-acetylglucosamine 2-epimerase"
/protein_id="AAM23927.1"
/db_xref="GI:20515649"
/translation="MKVITIGTRPQIFKAATVSRKLFNDMKIEIVHTGQHYDMM
SDIFFELEIPQPDYVYLGSGSHGQVTAHMLEKIEDVLIEKPDVLYIGDTNSTLA
GTIAASKLQIPVAHVAGLSRFRNKPPEINRIVADVSDLLFAPTKVAVNNLNEGI
ADRIYVGVDMVDAAFYFGRAGQSKILETLDLKPKVEYILSTIRAYNTDIPERLK
TIFNALCKIGKQDITVMPHLPPTKALLREGLYKVESEKLIIEPVYKIDMKVLEKNA
KLITDSGGVQKREAFYKVLVTLRETEWVELVGLGNVLPVPPDCEKFTIESIRKVL
NAPHGLDAPPIYDGGKAAEKIVKVLRYLL"
6228. .7175
/genes="WecB2"
/notes="Pfam match to entry Epimerase_2,
UDP-N-acetylglucosamine 2-epimerase, score 512.8, E-value
```

```
source          1.199450
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="4"
/clone="RP11-448G15"
BASE COUNT      54835 a 46671 c 45223 g 52721 t
ORIGIN
Query Match      4.2%; Score 49.6; DB 9; Length 199450;
Best Local Similarity 52.4%; Pred. No. 0.0012;
Matches 109; Conservative 0; Mismatches 99; Indels 0; Gaps 0;
Qy 74 TCTGTCGGGATTTAAATGATGACCGTCCCGTATTTAAACAATGTGATAAATPAC 133
Db 112372 TCTGTTGGAGCTAGTGCCACTGAGTCTTTGGGTGTACATCAGAATCACCTAGTC 112313
Qy 134 CGTTACGGGAAACCGCTGACAAATTCGGGCTGAAGAGATCCGGCTTATCTGTT 193
Db 112312 CCTGTCCCAACACACTGAATCAGAATGGCAGCGATGAGACTCTAGTGAATGTGAAGTT 112253
Qy 194 GCATTTCCCTTGTAGCTGACTAGCAGACACAATGATGTGCGCTTCTGTGTAATATC 253
Db 112252 GTATTTCCCTTGTGCTCATTTCTCTGTTGACTAATGATGTGAGGCAATCTGTCATTAGC 112193
Qy 254 AAACCGGTACTCAATATCTTCTGTGGCG 281.
Db 112192 ATACCAAGTCAATGATTTCTCTGGAG 112165

RESULT 31
AX433280/c      1135 bp DNA linear PAT 28-JUN-2002
LOCUS           Sequence 1695 from Patent WO0229113.
DEFINITION      AX433280
ACCESSION       AX433280
VERSION         AX433280.1 GI:21658084
KEYWORDS        Bacillus licheniformis.
SOURCE          Bacillus licheniformis.
ORGANISM        Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
REFERENCE       1
AUTHORS         Berka, R. and Clausen, I.G.
TITLE           Methods for monitoring multiple gene expression
JOURNAL         Patent: WO 0229113-A 1695 11-APR-2002;
Novozymes Biotech, Inc. (US) ; Novozymes A/S (DK)
FEATURES        Location/Qualifiers
source          1.1135
/organism="Bacillus licheniformis"
/db_xref="taxon:1402"
BASE COUNT      327 a 251 c 278 g 279 t
ORIGIN
Query Match      4.1%; Score 49; DB 6; Length 1135;
Best Local Similarity 50.2%; Pred. No. 0.0006;
Matches 121; Conservative 0; Mismatches 120; Indels 0; Gaps 0;
Qy 351 ACATCCCGCCAGGGTAAACAGCGTCCCTGTGCATATCTTCTGAATGATCATCAGGATCCCG 410
Db 977 ATATCTCCGACATCGACCAAGTAAACCGGATACCTCCGTTTTTGTATCATCCTCGGGAATCCCG 918
Qy 411 CCCGTCCTACCTGGGATAACGGGACGCGGAGACTGACGCTTCAGCCAGTACCATACCA 470
Db 917 CCGATATTGTGCGGATGCACGGAACCCGCGGCCATCGCTTCAGCAGCACTAAACCG 858
Qy 471 AACGCTTCATTTCCGAAGGATGACCAACCACTGGCAATCCGGTACCGGTAACGCT 530
Db 857 AAGCTTTCTTCTCGGACAAAGACAGCTTTAAATCGCTGATGCGGTAAAGCTTCTTACG 798
Qy 531 GGGAAAGGGCACTGCGCATTAACATATCCGCTCATNTCCAGGGTGTCTGTGCTGTA 590
Db 797 CTGTCTTGTGTCGAAGAAAGAAACGTCGTCTCTTTAATCCGAGCTGTCTGACAAGTGG 738
Qy 591 C 591

source          1.199450
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="4"
/clone="RP11-448G15"
BASE COUNT      54835 a 46671 c 45223 g 52721 t
ORIGIN
Query Match      4.1%; Score 48.6; DB 6; Length 783;
Best Local Similarity 49.8%; Pred. No. 0.00075;
Matches 123; Conservative 0; Mismatches 124; Indels 0; Gaps 0;
Qy 354 TCCCGCCAGGGTAAACAGCGTCCCTGTGCATATCTTCTGAATGATCATCAGGATCCCGCC 413
Db 631 TTCCCAAGGGGCATATAAATAACCACTTACTTGTCTTCAACAACACTTCCGGGATGCCACA 572
Qy 414 GTCTCACTGGCGATAACGGGACGCGGAGACTGACGCTTCAGCCAGTACCATACCAAC 473
Db 571 ATTTTGTGCGATCAGAGTACCGGCGATGCGATCGCTTCAAGCGCCACTAAACCAAG 512
Qy 474 GTTTCATTTTCGAAGGATGACACACACATCTCCGCTCATTTCCAGGTAAACGCTGGG 533
Db 511 CTTTCTTTTTCGTTAAACACAGCATTAATCGCTCATAGACACAAATTCAGCAACATGC 452
Qy 534 AAAAGGGCACCTGCCATTAAACATCTCCGCTCATTTCCAGGTGTCTGTCTGCTGACGC 593
Db 451 TTTTGTGTCCTAGCATAGGACATCGTTTGTCAAGCCGTTTCTTCCCAACAACATGCTC 392
Qy 594 AGACGTG 600
Db 391 GCAATTG 385

RESULT 33
AP000007/c      253505 bp DNA linear BCT 06-APR-2000
LOCUS           Pyrococcus horikoshii OT3 genomic DNA, 1485001-1738505 nt. position (7/7).
DEFINITION      AP000007
ACCESSION       AP000007 AB009465 AB009521 AB009522 AB009523 AB009524
AB009525 AB009526 AB009527 AB009528 AB009529 AB009530 AB009531
BA000001
VERSION         AP000007.1 GI:3236134
KEYWORDS        Pyrococcus horikoshii (strain:OT3) DNA.
SOURCE          Pyrococcus horikoshii
ORGANISM        Archaea; Euryarchaeota; Thermococci; Thermococcales;
Thermococcaceae; Pyrococcus.
REFERENCE       1 (sites)
AUTHORS         Kawarabayashi, Y., Sawada, M., Horikawa, H., Haikawa, Y., Hino, Y.,
Yamamoto, S., Sekine, M., Baba, S., Kosugi, H., Hosoyama, A., Nagai, Y.,
Sakai, M., Ogura, K., Otsuka, R., Nakazawa, H., Takamiya, M., Ohfuku, Y.,
Funahashi, T., Tanaka, T., Kudo, Y., Yamazaki, J., Kushiida, N.,
Oguchi, A., Aoki, K., Nakamura, Y., Robb, T.F., Horikoshi, K.,
Masuchi, Y., Shizuya, H. and Kikuchi, H.
TITLE           Complete sequence and gene organization of the genome of a
hyper-thermophilic archaeobacterium, Pyrococcus horikoshii OT3
```

JOURNAL DNA Res. 5 (2), 55-76 (1998)
 MEDLINE 98344137
 REFERENCE 2 (bases 1 to 253505)
 AUTHORS Tanaka, T., Kawarabayashi, Y. and Kikuchi, H.
 TITLE Direct Submission
 JOURNAL Submitted (11-JUN-1998) Yutaka Kawarabayashi, National Institute of Technology and Evaluation, Biotechnology Center; 2Chome 49-10 Nishihara, Shibuya-ku, Tokyo 151-0066, Japan
 (E-mail: genomeOT3@nite.go.jp, Tel: +81-3-3481-8951, Fax: +81-3-3481-8424)
 COMMENT On or before Mar 17, 1999 this sequence version replaced gi:3131896, gi:3131920, gi:3131974, gi:3131982, gi:3132032, gi:3132063, gi:3132118, gi:3132139, gi:3132184, gi:3132233, gi:3133142.
 AB009464-AB009465, AB009521-AB009531: submitted (10-DEC-1997) Kawarabayashi, Y. is officially affiliated with the National Institute of Bioscience and Human-Technology, Tsukuba, Ibaraki 305-0046, Japan.
 Robb, T. F. is at the Center of Marine Biotechnology, University of Maryland, Baltimore, MD, USA.
 Horikoshi, K. is at the Japan Marine Science and Technology Center, Yokosuka, Kanagawa 237-0061, Japan.
 Masuchi, Y. is at the University of Tokyo, Meguro, Tokyo 153-0041, Japan.
 Shizuya, H. is at the California Institute of Technology, Pasadena, CA, USA.
 The other authors are at the National Institute of Technology and Evaluation, 2-49-10 Nishihara, Shibuya, Tokyo 151-0066, Japan.
 All the sequence with length 100 codons or more between ATG or GTG and stop codon are defined as CDS
 Homology analysis is performed by Smith-Waterman algorithm against GenBank and GenPept release 103; EMBL release 52.0; SwissProt release 34.0; PIR-Protein release 54.0; and OWL release 29.5.
 E-mail address for comments and questions: genomeOT3@nite.go.jp
 Restriction map, ORF organization, sequence alignment and more information are available at W.W.W. site of Biotechnology Center, URL: <http://www.bio.nite.go.jp/>.
 FEATURES
 source Location/Qualifiers
 1. 253505
 /organism="Pyrococcus horikoshii"
 /strain="OT3"
 /db_xref="taxon:53953"
 /map="1485001-1738505 nt. position"
 BASE COUNT 74434 a 56827 c 47619 g 74625 t
 ORIGIN
 Query Match 4.1%; Score 48.2; DB 1; Length 253505;
 Best Local Similarity 59.0%; Pred. No. 0.0036;
 Matches 102; Conservative 0; Mismatches 68; Indels 3; Gaps 1;
 QY 352 CATCCCCCAGGGTAACAGCGTCCCTGTCACATCTCTCAATGACATCAGGATCCGC 411
 Db 118758 CATTTCCGGTGGAACTAAAGCCCTGCACTGTTTCTTTATAACCTCAGGTATACCAC 118699
 QY 412 CCCTCTCAGTCGGGTAACGGCGGACGCGGAGACTGACGCTTCAGCCAGTACCATAACAA 471
 Db 118698 CACGCTGTAGCTATATATGGGAACTCCGGAGGCGCATGGCTCAAGGATTACTATGCCGA 118639
 QY 472 ACCTGTCATTTTC---CGAAGGATGACACACCACTGGCAATCCGCTAGACC 521
 Db 118638 ACGCTTCGATGATATCGAAGGTAAGACGAACACGTCGGCCATCTCTGANAACC 118586
 RESULT 34
 SC2D46/c 35576 bp DNA linear BCT 12-MAY-2002
 LOCUS Streptomyces coelicolor cosmid 2D46.
 DEFINITION Streptomyces coelicolor cosmid 2D46.
 ACCESSION AL391406 AL645882
 VERSION AL391406.1 GI:9857143
 KEYWORDS AdaA-like; Arac-family; ATP/GTP-binding membrane; bldA; DNA-binding; GntR-family transcriptional regulator; hydrolase; integral; integrase; MarR-family; NLP/P60-family secreted; pgm; phosphoglycerate mutase; TetR-family; transferase; transport.

Streptomyces coelicolor A3(2).
 Streptomyces coelicolor A3(2)
 Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
 1 (bases 1 to 35576)
 Redenbach, M., Kieser, H. M., Denapaita, D., Eichner, A., Cullum, J.,
 Kinashi, H. and Hopwood, D. A.
 A set of ordered cosmids and a detailed genetic and physical map
 for the 8 Mb Streptomyces coelicolor A3(2) chromosome
 Mol. Microbiol. 21 (1), 77-96 (1996)
 97000351
 8843436
 2 (bases 1 to 35576)
 Brown, S. P. and Harris, D.
 Unpublished
 3 (bases 1 to 35576)
 Bentley, S. D., Parkhill, J., Barrell, B. G. and Rajandream, M. A.
 Direct Submission
 Submitted (03-AUG-2000) Streptomyces coelicolor sequencing project,
 Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge
 CB10 1SA E-mail: barrell@sanger.ac.uk Cosmids supplied by Prof.
 David A. Hopwood, (3) John Innes Centre, Norwich Research Park,
 Colney, Norwich, Norfolk NR4 7UH, UK
 Notes:
 Streptomyces coelicolor sequencing at The Sanger Centre is funded
 by the BBSRC and Beowulf Genomics
 Details of S. coelicolor sequencing at the Sanger Centre are
 available on the World Wide Web.
 (URL: <http://www.sanger.ac.uk/Projects/S.coelicolor/>) CDS are
 numbered using the following system eg SC787.01c. SC (S.
 coelicolor), 787 (cosmid name), .01 (first CDS), c (complementary
 strand).
 The more significant matches with motifs in the PROSITE database
 are also included but some of these may be fortuitous. The length
 in codons is given for each CDS.
 Usually the highest scoring match found by fasta -o is given for
 CDS which show significant similarity to other CDS in the database.
 The position of possible ribosome binding site sequences are given
 where these have been used to deduce the initiation codon. Gene
 prediction is based on positional base preference in codons using a
 specially developed Hidden Markov Model (Krogh et al., Nucleic
 Acids Research, 22(22):4768-4778(1994)) and the Frameplot program
 of Bibb et al., Gene 30:157-66(1984) as implemented at
<http://www.nih.go.jp/jun/cgi-bin/frameplot.pl>. CAUTION: We may not have predicted the
 correct initiation codon. Where possible we choose an initiation
 codon (atg, gtg, ttg or att) which is preceded by an upstream
 ribosome binding site sequence (optimally 5-13bp before the
 initiation codon). If this cannot be identified we choose the most
 upstream initiation codon.
 IMPORTANT: This sequence MAY NOT be the entire insert of the
 sequenced clone. It may be shorter because we only sequence
 overlapping sections once, or longer, because we arrange for a
 small overlap between neighbouring submissions. Cosmid 2D46 lies
 between and overlaps cosmids D66 and D8 on the AseI-D genomic
 restriction fragment.
 FEATURES
 source Location/Qualifiers
 1. 35576
 /organism="Streptomyces coelicolor A3(2)"
 /strain="A3(2)"
 /db_xref="taxon:100226"
 /clone="cosmid 2D46"
 1. 1200
 /note="Repeated degenerately at 1210..2414. This DNA
 repeat manifests as a duplication of a pairs of CDSs such
 that 2SCD46.01c is similar to 2SCD46.03c and 2SCD46.02 is
 similar to 2SCD46.04."
 1. 98
 /note="nominal overlap with Streptomyces coelicolor cosmid
 2SCD46
 Nominal overlap with Streptomyces coelicolor cosmid D66
 (EMBL:AL358692)"
 complement(142..147)
 repeat_unit
 misc_feature
 RBS


```
/protein_id="AAL82012.1"
/db_xref="GI:18894079"
/translation="MFRPLRLVSIILVGLVFLVLFQVYFVGQAILGGRRSRLLEATG
VGDYPIYLVIGVLMWVSPMEASVWGVRRLQRTGFESNWMYAPGILKIMVGLAIS
WMMSDILMLFALGVWIFDIKVAIYYAKAIPAIALAFVTFGLGFIYFAGLVMLK
NIGPLANIEFVILFSLGVFFPLSLPEIVRKISWIPLTHATSAVRKIFIGLSYSSV
VNEITAMLILLPVYWAISLAVKWAELKVRMMGYGY"
gene
CDS
/organism="Chlorobium tepidum TLS"
/strain="TLS"
/db_xref="taxon:194439"
complement(242..1099)
/genes="CT2209"
complement(242..1099)
/genes="CT2209"
/notes="identified by match to PFAM protein family HMM
PF00535"
/codon_start=1
/transl_table=11
/product="glycosyl transferase"
/protein_id="AAM73425.1"
/db_xref="GI:21648253"
/translation="MNGGTAVIVLGNWGAADTLACIASLAKVRQPAFTVILADNGST
DETVGLVQAFPEVEILEKRNIGLFAAGNNAAFSLRGRGDFRVVFLNNDTVVDPGFL
QPLLDLQRPWYIAAPKILYIMDDPGRIWAGVLESATGLIHTGIRDPDGPFRFTFP
EPVWYATGCCLAMRCRVFEVGGFDERFMYGDDVLSMKVREGLIVYQPARSLWH
RVSSSGGEMNLGKQLKSKAAMMLFAKHMIGGLVLYPLLLFRALLGLLRQPFWM
TSSQREEA"
gene
CDS
complement(1096..2121)
/genes="CT2210"
complement(1096..2121)
/genes="CT2210"
/notes="identified by match to PFAM protein family HMM
PF00534"
/codon_start=1
/transl_table=11
/product="glycosyl transferase"
/protein_id="AAM73426.1"
/db_xref="GI:21648254"
/translation="MFRRLYPRWLNGRTATEPGSDVPASPFTDLNPLSWITAA
RKLNRKIPDVLIAYWSGLAPICALMRRAAGLPTEVLLHNFTSHEPISGESMLKRL
VSSDGFITISOTVSELSAFAPAAKTLRLFLHPLYERQTPGPGKAARRSIGLPEDAP
VLIFGVYREYKGLDTLFEAMALVLRQESSARLVAGFETLSSRFREARRIGIDGA
VEFREGYVAGEVATLMAAADAVLVPYSATQSGIVPLALGHGVPVACDTGGIGNOV
EHRTGMLVREGEALADGILDFFRERELPLEEGISDFRRNSRREASQAATLE
TCSRSA"
gene
CDS
complement(2273..2539)
/genes="ptsH"
/notes="CT2211"
complement(2273..2539)
/genes="ptsH"
/notes="identified by match to PFAM protein family HMM
PF00381"
/codon_start=1
/transl_table=11
/product="phosphocarrier protein HPx"
/protein_id="AAM73427.1"
/db_xref="GI:21648255"
/translation="MIVQEVIRNSAGLHTRPAAAVVKLASRFSDFEIMDGLIENA
KSIIGVMSLAAPKGRSMVLKGEDEAAKHLIEFEEQFGEA"
gene
CDS
complement(2604..3071)
/genes="CT2212"
complement(2604..3071)
/genes="CT2212"
/notes="identified by match to PFAM protein family HMM
PF00583"
/codon_start=1
/transl_table=11
/product="acetyltransferase, GNAT family"
/protein_id="AAM73428.1"
/db_xref="GI:21648256"
/translation="MPAADTCVNRNARLADASAIRTEGYAGEGIMLKRSVENIIEHI
RQFFVADYQGVICCAIAFYTVKLAIRSLAVLEFRNKGIGRLIVERKAEVISEEG
VNEVFVLTLSGFFKRMGYKEIEYFPQKIWRDCTNCPRMACDEIAVVKTL"
gene
CDS
complement(3058..4065)
/genes="obg"
/notes="CT2213"
complement(3058..4065)
/genes="obg"
/notes="identified by match to PFAM protein family HMM
PF01018"
```

```
Query Match 4.0%; Score 47.4; DB 1; Length 10991;
Best Local Similarity 60.5%; Pred. No. 0.0033;
Matches 78; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

QY 350 CACATCCCGGAGGATACAGGTCCTCTGTCACATTCCTGTAATGACATCAGGGATCC 409
      ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 6404 CTCATCCAGGGGCACTAGAGGCCAGCTCTCTTTATGATTCAGGTATCC 6345
      ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 410 GCCCGTCTCAGCGGCAATACAGGCGACGCCGAGACGTGACGCTGACCGATCATTACC 469
      || || || || || || || || || || || || || || || || || || || ||
Db 6344 ACCAAGCTGTAGCTACAACTGGGACTCTGTGTCATTCGTCGACCACTATACC 6285
      ||||| |||||

QY 470 AAACGCTC 478
      ||||| |||||
Db 6284 AAATGCTC 6276

RESULT 37
AE012967
LOCUS Chlorobium tepidum TLS section 188 of 194 of the complete genome.
DEFINITION AE012967 AB008470
ACCESSION AE012967.1 GI:21648252
VERSION
KEYWORDS
SOURCE
ORGANISM Chlorobium tepidum TLS.
            Bacteria; Chlorobi; Chlorobia; Chlorobiales; Chlorobiaceae;
            Chlorobium.
REFERENCE 1 (bases 1 to 10148)
            Eisen, J.A., Nelson, K.E., Paulsen, I.T., Heidelberg, J.F., Wu, M.,
            Dodson, R.J., Deboy, R., Gwinn, M.L., Nelson, W.C., Haft, D.H.,
            Hickey, E.K., Peterson, J.D., Durkin, A.S., Kolonay, J.L., Yang, F.,
            Holt, I., Umayam, L.A., Mason, T., Brenner, M., Shea, T.P., Parksey, D.,
            Nierman, W.C., Feldblyum, T.V., Hansen, C.L., Craven, M.B., Radune, D.,
            Vamathevan, J., Khouri, H., White, O., Gruber, T.M., Ketchum, K.A.,
            Venter, J.C., Tettelin, H., Bryant, D.A. and Fraser, C.M.
            The complete genome sequence of Chlorobium tepidum TLS, a
            photosynthetic, anaerobic, green-sulfur Bacterium
            Proc. Natl. Acad. Sci. U.S.A. 99 (14), 9509-9514 (2002)
            12093901
REFERENCE 2 (bases 1 to 10148)
            Eisen, J.A., Nelson, K.E., Paulsen, I.T., Heidelberg, J.F., Wu, M.,
            Dodson, R.J., Deboy, R., Gwinn, M.L., Nelson, W.C., Haft, D.H.,
            Hickey, E.K., Peterson, J.D., Durkin, A.S., Kolonay, J.L., Yang, F.,
            Holt, I., Umayam, L.A., Mason, T., Brenner, M., Shea, T.P., Parksey, D.,
            Nierman, W.C., Feldblyum, T.V., Hansen, C.L., Craven, M.B., Radune, D.,
            Vamathevan, J., Khouri, H., White, O., Gruber, T.M., Ketchum, K.A.,
            Venter, J.C., Tettelin, H., Bryant, D.A. and Fraser, C.M.
            Direct Submission
            Submitted (30-APR-2002) The Institute for Genomic Research, 9712
            Medical Center Dr, Rockville, MD 20850, USA
            unpublished
            Location/Qualifiers
            1..10148
            source
```

```

/codon_start=1
/transl_table=11
/product="GTP-binding protein Obg"
/protein_id="AA073429.1"
/db_xref="GI:21648257"
/translation="MKFVDSAKISVKGAGDGGRCVSRREKVPKPGDGGDGRGGH
VLRANKQVTLDDFKYRSYTAGRGHGLGARKSGDKGKDVIIIGVPGTVVVRNVEVG
EVLCDMVEQGEQIMAKGGGCGWGNQHPATARAPRAQPGCEGEVELEMELEKLM
DVLGVFPNAGSTLLSVLSARPKIADYPTFTLVPNLGIYVYEDYKSPVADIPGII
EGAEGRGIGIQFRIERTKILLIMVPSNEDIAEAYATLKELEKEDFSPSLRAPRL
VWTRMDIAPDETPELEKGVKVLTAISSVAGGLKALKDELWRQVSLQNSPSEHAG
S"
4135..4422
/gene="xseB"
/note="CT2214"
4135..4422
/gene="xseB"
/note="identified by match to PFAM protein family HMM
PF02609"
/codon_start=1
/transl_table=11
/product="exodeoxyribonuclease, small subunit"
/protein_id="AA073430.1"
/db_xref="GI:21648258"
/translation="MPASKSRTSAYPTIEELIQRLEETRNIENPDTGLNSIALYE
EGMSLAPECKRLETRKLEETINPAETARPAKPNAPESPRMNDLFGTES"
complement(4437..5864)
/gene="gatB"
/note="CT2215"
complement(4437..5864)
/gene="gatB"
/note="identified by match to PFAM protein family HMM
PF01162"
/codon_start=1
/transl_table=11
/product="glutamy1-tRNA(Gln) amidotransferase subunit B"
/protein_id="AA073431.1"
/db_xref="GI:21648259"
/translation="MVEIVVGLVHVCQLNTESKAFKCGGSAKFGKPNANTVPCVIAL
PGALPVLNARVVEDAVKLGATNCTIARHSILARKNFYFDLPKGVQISQYEPIGCE
GVTHIDLEEGKDVLRVRIHIEDACKSIHIDGDTYIDVNRGCVPLLEIVSPDMRT
PKRASAYLQKRLQIVRYLIGSDGNMEGSLRCDANVSRVPGATEYGTETIKMNSF
RNVERLEYAKRHQIEVEGGGTIVQETRLWDADKLETRSMRGKERAHDYRFPDDVL
VPLVDDGMTRHOEELEPEPDRARFVSEFGIPAYDAGVITVDRELADYFESTVPL
SGDAKSNVVMGVEVMTLKEKYLIDHKFAISPERLGLIKLINAGASINTAKQVE
IMQODEATAIVEREGLAQVSDRGAIEAIREILEANOKLEQYRSKTKLQFGFVFG
CQMKMGKANPKMNDILRSLDA"
complement(5910..6491)
/gene="CT2216"
complement(5910..6491)
/gene="CT2216"
/note="identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AA073432.1"
/db_xref="GI:21648260"
/translation="MILLSNEKVMAMKRVHLMRALIALAATVLMFGTARAGWDPAD
EVNARTIMFRKADPSLRFKSGYAVFSYIKGFMIVGGHGKGVVFDQGRPI
GHAVTFMNGVPGOLGGSFSELIIFKDRALADFTKGNYSQAQFAVAVRAGMATNT
DYSNGVAFAMPNAGLMGELTVGAQFTFRQYR"
6475..6603
/gene="CT2217"
6475..6603
/gene="CT2217"
/note="identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AA073433.1"
/db_xref="GI:21648261"
/translation="MDNKIRPAKSNATPAYNVRFSSLKTRIASERKKPKNQLVGR"
6681..6824
/gene="CT2218"
6681..6824
/note="identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AA073434.1"
/db_xref="GI:21648262"
/translation="MQGHAISIQPRCVITPMLYAFKELATQCPCPAYSAFDSIYTYFP
WRY"
complement(6837..7355)
/gene="CT2219"
complement(6837..7355)
/gene="CT2219"
/note="identified by match to PFAM protein family HMM
PF01661"
/codon_start=1
/transl_table=11

```

```

CDS
/gene="CT2218"
6681..6824
/note="identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AA073434.1"
/db_xref="GI:21648262"
/translation="MQGHAISIQPRCVITPMLYAFKELATQCPCPAYSAFDSIYTYFP
WRY"
complement(6837..7355)
/gene="CT2219"
complement(6837..7355)
/gene="CT2219"
/note="identified by match to PFAM protein family HMM
PF01661"
/codon_start=1
/transl_table=11

gene
Query Match 4.0%; Score 47.2; DB 1; Length 10148;
Best Local Similarity 50.4%; Pred. No. 0.0038;
Matches 115; Conservative 0; Mismatches 113; Indels 0; Gaps 0;

QY 281 GCTGCTGCCATCATCCGGAAGCGTTCGGTGGGATAAAAAATCGCGAGTGGCGGGT 340
Db 1183 GATGCCCTCTCCAGCGGCGGCTCTCGCTCGCGGAAAAAGTCGAGAATCCGTCAGC 1242

QY 341 CCATGCAGACATCCCCACGCGGTAACAGCGTCCCTGCATCTTCTGAATGACATC 400
Db 1243 CAGTCCCTCGGGCCCTCTCTCGCGAGCAGCCGCGGTAGCTCGCTCGACCTGATT 1302

QY 401 AGGGATCCGCCCGCTCTCACTGGCGATAACGGGACACGGCGAGACTGACGCTTCAGCCAG 460
Db 1303 GCGAGGCGCGGGTGTCTCGAGCGCATCACCGGACGCGGTGGCGAGCGAGTGGCAC 1362

QY 461 TACCATACCAACCGTTTCATTTTCGAAGGATGACACACACACTGGC 508
Db 1363 GATGCCCGACTCGTGGCGGACGCGTAGGCGAGCACCACCGCATCGGC 1410

RESULT 38
LOCUS AX437696 783 bp DNA linear PAT 28-JUN-2002
DEFINITION Sequence 6111 from Patent WO0229113.
ACCESSION AX437696
VERSION AX437696.1 GI:21662504
KEYWORDS Bacillus clausii.
SOURCE Bacillus clausii.
ORGANISM Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
REFERENCE 1
AUTHORS Berka,R. and Clausen,I.G.
TITLE Methods for monitoring multiple gene expression
JOURNAL Patent: WO 0229113-A 6111 11-APR-2002;
FEATURES
source
1. .783
/organism="Bacillus clausii"
/db_xref="taxon:79880"
BASE COUNT 224 a 150 c 204 g 203 t 2 others
ORIGIN

Query Match 3.9%; Score 46.2; DB 6; Length 783;
Best Local Similarity 59.5%; Pred. No. 0.0046;
Matches 78; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

QY 354 TCCCCACCGGTAAACAGCGTCCCTGTCACATTTCTGTGATCATCAGGATCCGCC 413
Db 633 TTCCCAAGGGGCGCAATAATACCAAGTACTTCGTCTTCAACACTTCGGGATGCCACCA 574
QY 414 GTCTACTCGCGGATACGGGACGCGGAGACTGACGCTTACGCCAGTACCAATACCAAC 473

```

Dd 573 ATTTTGTGGGATACAGGTACCCTCCCATCGCTTCGTTCAAGCGGCCTAAACCAAG 514

Oy 474 GCTTCATTTC 484
|||||

Dd 513 CTTTCATTTC 503

RESULT 39
LOCUS AE004745
DEFINITION Pseudomonas aeruginosa PA01, section 306 of the complete genome. BCT 30-AUG-2000

ACCESSION AE004745 AE004091
VERSION AE004745.1 GI:9949336

KEYWORDS
SOURCE Pseudomonas aeruginosa.
ORGANISM Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
Pseudomonas.

REFERENCE 1 (bases 1 to 11884)
AUTHORS Stover,C.K., Pham,X.-Q.T., Erwin,A.L., Mizoguchi,S.D., Warrenner,P., Hickey,M.J., Brinkman,F.S., Huftnagle,W.O., Kowalik,D.J., Lagrou,M., Garber,R.L., Goltry,L., Tolentino,E., Westbrook-Wadman,S., Yuan,Y., Brody,L.L., Coulter,S.N., Folger,K.R., Kas,A., Larbig,K., Lim,R., Smith,K., Spencer,D., Wong,G.K., Wu,Z. and Paulsen,I.T.
TITLE Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathogen
JOURNAL Nature 406 (6799), 959-964 (2000)
MEDLINE 20437337
PubMed 10984043

REFERENCE 2 (bases 1 to 11884)
AUTHORS Stover,C.K., Pham,X.-Q.T., Erwin,A.L., Mizoguchi,S.D., Warrenner,P., Hickey,M.J., Brinkman,F.S., Huftnagle,W.O., Kowalik,D.J., Lagrou,M., Garber,R.L., Goltry,L., Tolentino,E., Westbrook-Wadman,S., Yuan,Y., Brody,L.L., Coulter,S.N., Folger,K.R., Kas,A., Larbig,K., Lim,R.M., Smith,K.A., Spencer,D.H., Wong,G.K.-S., Wu,Z., Paulsen,I.T., Reizer,J., Sailer,M.H., Hancock,R.E.W., Lory,S. and Olson,M.V.
TITLE Direct Submission
JOURNAL Submitted (16-MAY-2000) Department of Medicine and Genetics,
University of Washington Genome Center, University Of Washington,
Box 352145, Seattle, WA 98195, USA

FEATURES
source location/Qualifiers
1..11884
/organism="Pseudomonas aeruginosa"
/strain="PA01"
/db_xref="taxon:287"
102..419
/gene="PA3216"
102..419
/gene="PA3216"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAG06604.1"
/db_xref="GI:9949337"
translation="MSSPDDERFHSFAEFPYILQHSNVCRRLLHYVCSLLVLAVLH
HAVASQQWLWLLALPLGYCYGANWGCHVFYEKNRPATFYPLPSFMGDNVLMKDAFTQY
IRF"

636..2027
/gene="PA3217"
636..2027
/gene="PA3217"
/codon_start=1
/transl_table=11
/product="probable adenylyate cyclase"
/protein_id="AAG06605.1"
/db_xref="GI:9949338"
translation="MKPTLPDRSPTRSRVSAMREYSVLYAIACGASTACTTYIQY
SYGLMWVPYALLPHGLAYHLGCRQPRDKRVTRALLADVAHVCGMLGFGSVV
SLMFLLVLSFTALVIGLRGMALLVSASSALLVALVAPLLGTPPEVAASVLI
CGLYCITAFTHGQOGRLAQRVEAREAKRLARLNARYLSPOVWEMTFSGKK
VRLETRBKTKVTFFDSINGFTFLSEEEALEALTOLLNNYNEMSIALKYIGDTIDKF

636..2027
/gene="PA3217"
636..2027
/gene="PA3217"
/codon_start=1
/transl_table=11
/product="probable adenylyate cyclase"
/protein_id="AAG06605.1"
/db_xref="GI:9949338"
translation="MKPTLPDRSPTRSRVSAMREYSVLYAIACGASTACTTYIQY
SYGLMWVPYALLPHGLAYHLGCRQPRDKRVTRALLADVAHVCGMLGFGSVV
SLMFLLVLSFTALVIGLRGMALLVSASSALLVALVAPLLGTPPEVAASVLI
CGLYCITAFTHGQOGRLAQRVEAREAKRLARLNARYLSPOVWEMTFSGKK
VRLETRBKTKVTFFDSINGFTFLSEEEALEALTOLLNNYNEMSIALKYIGDTIDKF

Qy 426 ATACGGGACGCGGAGACTGACCTTCAGCGAGTACCATACCAACCGCTTCATTTC 485
Db 3189 ACACGAGGACACACACGACCTGCTTCAGCAAGAACCATCCGGAAGCCCTCCCTGGTG 3248
Qy 486 GAAGGATGACACACACGACCTGCGCAATCCGG 515
Db 3249 GATGGAAGGACACGACCTGCGCCCGCCCG 3278

RESULT 41
AF498404/c 13401 bp DNA linear BCT 13-JUN-2002
DEFINITION Pseudomonas aeruginosa serotype 013 putative O-antigen biosynthesis
gene cluster, partial sequence.

ACCESSION AF498404 AC104723
VERSION AF498404.1 GI:20559835
KEYWORDS
SOURCE
ORGANISM Pseudomonas aeruginosa.

Pseudomonas aeruginosa
Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
Pseudomonas.

REFERENCE 1 (bases 1 to 13401)
Raymond,C.K., Sims,E.H., Kas,A., Spencer,D.H., Kutayavin,T.V.,
Ivey,R.G., Zhou,Y., Kaul,R., Clendinning,J.B. and Olson,M.V.
Genetic variation at the O-antigen biosynthetic locus in
Pseudomonas aeruginosa
J. Bacteriol. 184 (13), 3614-3622 (2002)

REFERENCE 2 (bases 1 to 13401)
Raymond,C.K., Sims,E.H., Kas,A., Spencer,D.H., Kutayavin,T.V.,
Ivey,R.G., Zhou,Y., Kaul,R., Clendinning,J.B. and Olson,M.V.
Direct Submission
Submitted (04-APR-2002) Genome Center, University of Washington,
Box 352145, Seattle, WA 98105-2145, USA
On May 13, 2002 this sequence version replaced gi:17975252.
Location/Qualifiers

1. 13401
/organism="Pseudomonas aeruginosa"
/serotype="013"
/db_xref="taxon:287"
1. 104
/note="ORF_1: C-terminal coding region of rpsA"
/codon_start=3
/transl_table=11
/product="RpsA"
/protein_id="AA027608.1"
/db_xref="GI:20559836"
/translation="EKDAMKELKQEVESAGPTTIGDLIRAQMENQG"
241..525
/note="ORF_2: hmd/ihfB"
/codon_start=1
/transl_table=11
/product="Hmd"
/protein_id="AA027609.1"
/db_xref="GI:20559837"
/translation="MTKSLERIVTHOGLSAKQVELAKTLMLEQMSQALATGDRIE
IRGFGFSHLHAPRVGRNPKTGESVRLDGKFPKPKGKELDRVNEP"
597..815
/note="ORF_3: hypothetical orfA"
/codon_start=1
/transl_table=11
/protein_id="AA027610.1"
/db_xref="GI:20559838"
/translation="WVALTVIFVLENQSGVNLSTLWRTPTGPLSLFVSLAFVLGGL
GGVLSGMLRLIGSVSKGHLVPRKAM"
1167..2216
/note="ORF_4: wzz; similar to chain length determinant
protein"
/codon_start=1
/transl_table=11
/product="Wzz"
/protein_id="AA027611.1"

REFERENCE 3 (bases 1 to 13401)
Raymond,C.K., Sims,E.H., Kas,A., Spencer,D.H., Kutayavin,T.V.,
Ivey,R.G., Zhou,Y., Kaul,R., Clendinning,J.B. and Olson,M.V.
Direct Submission
Submitted (04-APR-2002) Genome Center, University of Washington,
Box 352145, Seattle, WA 98105-2145, USA
On May 13, 2002 this sequence version replaced gi:17975252.
Location/Qualifiers

1. 13401
/organism="Pseudomonas aeruginosa"
/serotype="013"
/db_xref="taxon:287"
1. 104
/note="ORF_1: C-terminal coding region of rpsA"
/codon_start=3
/transl_table=11
/product="RpsA"
/protein_id="AA027608.1"
/db_xref="GI:20559836"
/translation="EKDAMKELKQEVESAGPTTIGDLIRAQMENQG"
241..525
/note="ORF_2: hmd/ihfB"
/codon_start=1
/transl_table=11
/product="Hmd"
/protein_id="AA027609.1"
/db_xref="GI:20559837"
/translation="MTKSLERIVTHOGLSAKQVELAKTLMLEQMSQALATGDRIE
IRGFGFSHLHAPRVGRNPKTGESVRLDGKFPKPKGKELDRVNEP"
597..815
/note="ORF_3: hypothetical orfA"
/codon_start=1
/transl_table=11
/protein_id="AA027610.1"
/db_xref="GI:20559838"
/translation="WVALTVIFVLENQSGVNLSTLWRTPTGPLSLFVSLAFVLGGL
GGVLSGMLRLIGSVSKGHLVPRKAM"
1167..2216
/note="ORF_4: wzz; similar to chain length determinant
protein"
/codon_start=1
/transl_table=11
/product="Wzz"
/protein_id="AA027611.1"

REFERENCE 4 (bases 1 to 13401)
Raymond,C.K., Sims,E.H., Kas,A., Spencer,D.H., Kutayavin,T.V.,
Ivey,R.G., Zhou,Y., Kaul,R., Clendinning,J.B. and Olson,M.V.
Direct Submission
Submitted (04-APR-2002) Genome Center, University of Washington,
Box 352145, Seattle, WA 98105-2145, USA
On May 13, 2002 this sequence version replaced gi:17975252.
Location/Qualifiers

1. 13401
/organism="Pseudomonas aeruginosa"
/serotype="013"
/db_xref="taxon:287"
1. 104
/note="ORF_1: C-terminal coding region of rpsA"
/codon_start=3
/transl_table=11
/product="RpsA"
/protein_id="AA027608.1"
/db_xref="GI:20559836"
/translation="EKDAMKELKQEVESAGPTTIGDLIRAQMENQG"
241..525
/note="ORF_2: hmd/ihfB"
/codon_start=1
/transl_table=11
/product="Hmd"
/protein_id="AA027609.1"
/db_xref="GI:20559837"
/translation="MTKSLERIVTHOGLSAKQVELAKTLMLEQMSQALATGDRIE
IRGFGFSHLHAPRVGRNPKTGESVRLDGKFPKPKGKELDRVNEP"
597..815
/note="ORF_3: hypothetical orfA"
/codon_start=1
/transl_table=11
/protein_id="AA027610.1"
/db_xref="GI:20559838"
/translation="WVALTVIFVLENQSGVNLSTLWRTPTGPLSLFVSLAFVLGGL
GGVLSGMLRLIGSVSKGHLVPRKAM"
1167..2216
/note="ORF_4: wzz; similar to chain length determinant
protein"
/codon_start=1
/transl_table=11
/product="Wzz"
/protein_id="AA027611.1"

REFERENCE 5 (bases 1 to 13401)
Raymond,C.K., Sims,E.H., Kas,A., Spencer,D.H., Kutayavin,T.V.,
Ivey,R.G., Zhou,Y., Kaul,R., Clendinning,J.B. and Olson,M.V.
Direct Submission
Submitted (04-APR-2002) Genome Center, University of Washington,
Box 352145, Seattle, WA 98105-2145, USA
On May 13, 2002 this sequence version replaced gi:17975252.
Location/Qualifiers

1. 13401
/organism="Pseudomonas aeruginosa"
/serotype="013"
/db_xref="taxon:287"
1. 104
/note="ORF_1: C-terminal coding region of rpsA"
/codon_start=3
/transl_table=11
/product="RpsA"
/protein_id="AA027608.1"
/db_xref="GI:20559836"
/translation="EKDAMKELKQEVESAGPTTIGDLIRAQMENQG"
241..525
/note="ORF_2: hmd/ihfB"
/codon_start=1
/transl_table=11
/product="Hmd"
/protein_id="AA027609.1"
/db_xref="GI:20559837"
/translation="MTKSLERIVTHOGLSAKQVELAKTLMLEQMSQALATGDRIE
IRGFGFSHLHAPRVGRNPKTGESVRLDGKFPKPKGKELDRVNEP"
597..815
/note="ORF_3: hypothetical orfA"
/codon_start=1
/transl_table=11
/protein_id="AA027610.1"
/db_xref="GI:20559838"
/translation="WVALTVIFVLENQSGVNLSTLWRTPTGPLSLFVSLAFVLGGL
GGVLSGMLRLIGSVSKGHLVPRKAM"
1167..2216
/note="ORF_4: wzz; similar to chain length determinant
protein"
/codon_start=1
/transl_table=11
/product="Wzz"
/protein_id="AA027611.1"

REFERENCE 6 (bases 1 to 13401)
Raymond,C.K., Sims,E.H., Kas,A., Spencer,D.H., Kutayavin,T.V.,
Ivey,R.G., Zhou,Y., Kaul,R., Clendinning,J.B. and Olson,M.V.
Direct Submission
Submitted (04-APR-2002) Genome Center, University of Washington,
Box 352145, Seattle, WA 98105-2145, USA
On May 13, 2002 this sequence version replaced gi:17975252.
Location/Qualifiers

1. 13401
/organism="Pseudomonas aeruginosa"
/serotype="013"
/db_xref="taxon:287"
1. 104
/note="ORF_1: C-terminal coding region of rpsA"
/codon_start=3
/transl_table=11
/product="RpsA"
/protein_id="AA027608.1"
/db_xref="GI:20559836"
/translation="EKDAMKELKQEVESAGPTTIGDLIRAQMENQG"
241..525
/note="ORF_2: hmd/ihfB"
/codon_start=1
/transl_table=11
/product="Hmd"
/protein_id="AA027609.1"
/db_xref="GI:20559837"
/translation="MTKSLERIVTHOGLSAKQVELAKTLMLEQMSQALATGDRIE
IRGFGFSHLHAPRVGRNPKTGESVRLDGKFPKPKGKELDRVNEP"
597..815
/note="ORF_3: hypothetical orfA"
/codon_start=1
/transl_table=11
/protein_id="AA027610.1"
/db_xref="GI:20559838"
/translation="WVALTVIFVLENQSGVNLSTLWRTPTGPLSLFVSLAFVLGGL
GGVLSGMLRLIGSVSKGHLVPRKAM"
1167..2216
/note="ORF_4: wzz; similar to chain length determinant
protein"
/codon_start=1
/transl_table=11
/product="Wzz"
/protein_id="AA027611.1"

/db_xref="GI:20559839"
/translation="MTEETFFQVAKNEEDFLGLRLMAEKWLIACITILFTSLVG
TYATLSPVYEAIRAVPPSPSAGNLAENGLATTSVNDVYSVFLRNLAETRR
NRFREYVLPSELSERSGSKDRYLSNLSERIEQRPQSDREERYTVLFEKHDQQA
DWGRYIKLAAEGSKMDQAVREFEVKSNIOQDILLDRADAKAREDIRNPKKA
YQIAESKLENPPLIGGMDQQLSSIMEGALMYNRGTAKALRAELREERTSDPPFP
ALRNQEQAMYSTIKLDVNVKAVFRQDQVAEVPDPVRPRKKLLLVIGALSGLVWG
MIATVMAVRRSRK"
2303..3568
/note="ORF_5: similar to UDP-glucose/GDP-mannose
dehydrogenase"
/codon_start=1
/transl_table=11
/protein_id="AA027612.1"
/db_xref="GI:20559840"
/translation="MDIKAAVGLVGLPLAVEFGKRSVTGFDINHSTRDLQKY
DSTLEVEKEELSAVHLKFTSTLSELOECNEFFIVTPTIDEHKQDPLPLVKAESI
AKVLKNDIVYESTVPGATEEVCVPLERESGLRNSDFVGYSPERINPGDKEHR
VSTIKKVTSGSTPEAEIVDSLYREIITAGTHKAESIKVAEAKVIENTQDNLIALI
NELAIIFNKLEIDTESVLAAGTWNFLPFRGLVGGHCGVDYPIYTHKAQSTGYHP
EIIAGRLNDMGAYVVSQVKAHLKRRIHVDGARVLVGLTFKENCPLDNRNKVD
IVRLAEYNIADVDPWVSVEEAQHEVGLTPISAPVEGNYDAVVLAVAHNEFKELGA
DKIRAFGKLESVIYDLKILDKSDSLR"
3578..4603
/note="ORF_6: similar to NAD dependent
epimerase/dehydratase family"
/codon_start=1
/transl_table=11
/protein_id="AA027613.1"
/db_xref="GI:20559841"
/translation="MNAVAKALSRETTSSWLVTVGAGTGSNLELLKFNQKVVG
LDSTGHMNLDEVLGVEERWNEFKIQGDIRDIDCDRRAVEGYDVLHQALGS
VPSIADPTTNSVNIISGFLNIVASKEAGVKSITVYAASSYTGHDGPKRVEVIGK
PLSPVATKYVVELYADVFCYCGNLIGRYFNFGKRODPNGAYAAVLPKWTSMI
RGPVYINGDGETSRDCCFINNVIKANLAAFCETEAKNOVYNVAGRTTGLQDP
SLKVELSBNGINYEQGPYRFRAGDVHRSQAETKKAQVLYGYNPEYDIQAGIAMP
WYISFVK"
4665..5954
/note="ORF_7: similar to Polysaccharide biosynthesis
protein; potential multiple membrane spanning domains."
/codon_start=1
/transl_table=11
/protein_id="AA027614.1"
/db_xref="GI:20559842"
/translation="MNPKNIFWNLGLPLIIAALVPHLIATVVERFGFLASWG
LIGYALDILGIGRAVTKLSSIRGSEEMNSYITMTRITAVIGLGFSLIIIFG
LIGGEFFSHKDSVSELEFVLLALFALQOASATYRGVNEAYLNFKNLRIPL
GATNFGPFLSVYAKOLHYLVTLVLSRALAFFIRLHAGVLEKTELVERSCYDR
QAVELFRGGHWTVSSINPELVQSDRFFIGVLLSAAAVSYVPIEYITQSMILVGA
VSTVAFPSISNLRTSFAELCEFNKLLRVLLIMGGMLCLAFLLPILKLMWGDYI
GDSISVGRILCLGVFFNALGAMFYSLHNAKVKETAILHSIELPILILILIPR
HGIVGAAVWSLRLTVDITVLAVLSYFGWRVFPANC"
6001..7176
/note="ORF_8: similar to Sodium/hydrogen exchanger family;
potential multiple membrane spanning domains."
/codon_start=1
/transl_table=11
/protein_id="AA027615.1"
/db_xref="GI:20559843"
/translation="MVETALIGIKLILLLALLFTLISSLARADNLVYLLAGLFS
GVGLWSYIGELRNAGALQVLTVMVYVPLVLLAVSYRQSEALGKLPFYAALLIA
VLIDIFFVFNGLFSYNPYVSEMAAYGDAVVDNAGRIKFTLPNVSSLIFLIPYVS
YLLCKFSLGLFLALLVAVSGRALLVAFAPFLAYFLTMGADRQWQVIVSR
VLLFVFMGLFLIFYNTPYVENLFLSINFSNDNSNRYOFEALMKGYQOQPF
FGMGAADYADIRSEMPWAYELFVSLVFOYGILGFLYALGVLELLVFLVLRVKI
GRSPFECVLSGMLAFMATATNPYLAKFDYMWVIFIPVALLNYSVSRGDKGSNI"
7157..8092
/note="ORF_9: similar to Glycosyl transferase"
/codon_start=1
/transl_table=11
/protein_id="AA027616.1"
/db_xref="GI:20559844"
/translation="NESFIIDVVVWNWAGLQLRACVDSIFKSSGLIRKIVVIDNGS

promoter 1684..1715
/note="factor Sigma70; predicted +1 start at 4141533"
gene 1763..4060
/gene="pfid"
/note="b3951"
1763..4060
/gene="pfid"
/EC number="2.3.1.54"
/function="enzyme; Energy metabolism, carbon: Anaerobic
respiration"
/note="o765; 100 pct identical amino acid sequence and
equal length to PFLD_ECOLI SW: P32674"
/codon_start=1
/transl_table=11
/product="formate acetyltransferase 2"
/protein_id="AAC76933.1"
/db_xref="GI:1790388"
/translation="MTNRSRLKLTALFANTREISLERALLYTASHRQTEGEPIVLRR
KATAYILHEVEISIRDEELTAGNRTVYKAGIMSPENDPYLLKELDQPTTPQDRFA
ISEDKRIYREILFPYWEKRMKDFNGQMTDEVKAATNTQIFSIQTDKGGHIIID
YPLLNGHGLVAOQHOCHQCOQPHNFYQALLLEASOKHILRYVAELAEYMAANCT
DAORRELLTAEISRNQAKHPQTFWQACOLPYWYNNIILOVESNASSISLGRFOQYM
LPYQISLQGEAFAELKELLESILWVKNDIVLLRSTSSARFPAGPTGYTALLGLT
ENGSANVLSFCLDAYQSVQPOPNLGVRTNALIDTFELMKTAETIRFGTGPOIF
NDEVVPAFLNRGVSLDARDYVSGCVLSIPGRTYGLDILAMFNLLKVMETCLHEN
EGNAALTYEGLLQIRAKISHYITLMVEGSDICDIGHRDWAPVPLLSFISDLCKEGR
DITDGGARYNSFGVGGIGIANLSDSLHAKGMVFEOQLRSEDELLSVKANPATEGE
KVARLNRKPEKGNIDVDNISALLRHCKEVEKYONPRGGYTPGCVYVSAHPV
LGSVGTAPGRAGQLADGGLSPMLQDQAGPTAVLKSVDKLDNTLLSNGTLNVK
FTPATLEGAGLRADFLRAFTQLKLIHQIFNVNADTLREAAQRPQDIAGLVLRVA
GYSAFFVELSKETODDIIRTAHQL"
3292..3329
/note="REP (repetitive extragenic palindromic) element;
contains 1 REP sequence"
4026..4904
/gene="pfic"
/note="b3952"
4026..4904
/gene="pfic"
/EC number="1.97.1.4"
/function="putative enzyme; Energy metabolism, carbon:
Anaerobic respiration"
/note="o292; 100 pct identical to PFLC_ECOLI SW: P32675"
/codon_start=1
/transl_table=11
/product="probable pyruvate formate lyase activating
enzyme 2"
/protein_id="AAC76934.1"
/db_xref="GI:1790389"
/translation="MTSSAGORISCNVETRRDDVARIFNIQRYSLNDGEGITVVVFF
KGCPLCPWCANPESISGKTQTVRRREACLHCAKLRDADPCPSGAFERIGRDISLDA
LERVMKDDIFFFTSGGVTLSGGEVLMQAFATRLQRLRLWGVSCATETAGDAPAS
KLPLAKICDEVLDKIMDATQARDVKNMLPVLENRLRLVSEGVNVPRLPLIPG
FTLSRENQALDVLPLNIRQHLPLFPFHQYGEPKYLLGKTWSMEVAPSSADVAT
ARENAERAGLQVTVGG"
4730..4760
/gene="pfic"
/note="factor Sigma70; predicted +1 start at 4144578"
4864..4894
/gene="pfic"
/note="factor Sigma70; predicted +1 start at 4144712"
4906..5247
/gene="frwd"
/note="b3953"
4906..5247
/gene="frwd"
/function="enzyme; Degradation of small molecules: Carbon
compounds"
/note="o113; 100 pct identical amino acid sequence and
equal length to PTWX_ECOLI SW: P32676"
/codon_start=1
/transl_table=11
/product="PTS system fructose-like IIB component 2"

/protein_id="AAC76935.1"
/db_xref="GI:1790390"
/translation="MAYLVAVTACVSGVHAHTYMAAERLEKLLCKLVSTETOGALG
TENRLADEDIRADVALITDIELAGAEFHCERYVQCSIYAFLEPQVMSAVRKVL
SAPQOHLILE"
complement(5234..6085)
/gene="yijO"
/note="b3954"
complement(5234..6085)
/gene="yijO"

Query Match 3.8%; Score 45.4; DB 1; Length 10940;
Best Local Similarity 63.1%; Pred. No. 0.015;
Matches 70; Conservative 0; Mismatches 41; Indels 0; Gaps 0;
QY 11 TCGAAAAAGCCAAATAAAAAATGCCATCCCGCGCTCCAGCTGAAGTAGGCC 70
DB 7949 TAGAGAAAAACCAAAAGTAGAGCAGTCCCAACCGAGGCTTCCAGCTAAAAAGAGGTT 8008
QY 71 TGTTCGTCCGGTATTAAATGCATTCACCGTCCCGCTTATTAAACAATGT 121
DB 8009 TAGCCTGGACTTCTCTGGAATCCTAGTAAGAACCCTGCTTGAAAAAATAT 8059

RESULT 44
LMFLCHR32_19
WPCOMMENT
Sequence split into 28 fragments Locus LMFLCHR32 Accession AL499622

Fragment Name	Begin	End
LMFLCHR32_00	1	110000
LMFLCHR32_01	100001	210000
LMFLCHR32_02	200001	310000
LMFLCHR32_03	300001	410000
LMFLCHR32_04	400001	510000
LMFLCHR32_05	500001	610000
LMFLCHR32_06	600001	710000
LMFLCHR32_07	700001	810000
LMFLCHR32_08	800001	910000
LMFLCHR32_09	900001	1010000
LMFLCHR32_10	1000001	1110000
LMFLCHR32_11	1100001	1210000
LMFLCHR32_12	1200001	1310000
LMFLCHR32_13	1300001	1410000
LMFLCHR32_14	1400001	1510000
LMFLCHR32_15	1500001	1610000
LMFLCHR32_16	1600001	1710000
LMFLCHR32_17	1700001	1810000
LMFLCHR32_18	1800001	1910000
LMFLCHR32_19	1900001	2010000
LMFLCHR32_20	2000001	2110000
LMFLCHR32_21	2100001	2210000
LMFLCHR32_22	2200001	2310000
LMFLCHR32_23	2300001	2410000
LMFLCHR32_24	2400001	2510000
LMFLCHR32_25	2500001	2610000
LMFLCHR32_26	2600001	2710000
LMFLCHR32_27	2700001	272709

Continuation (20 of 28) of LMFLCHR32 from base 1900001 (AL499622 Leishmania major chr

Query Match 3.8%; Score 45.4; DB 2; Length 110000;
Best Local Similarity 51.2%; Pred. No. 0.025;
Matches 106; Conservative 0; Mismatches 101; Indels 0; Gaps 0;
QY 337 CGGTCCATCAGACATCCCCACGGGTAAACAGGTCCTCTCACATTTCTTGAATGA 396
DB 94357 CGGCCAGTCGGCGCGGATGCTGTTGGCGGCACCAAGTAGCGGTGCGCGCTTCAGCACGG 94416
QY 397 CATCAGGGATCCGGCCGCTCTCCTGCGGATACCGGACGCGCGAGACTGACGCTTACG 456
DB 94417 TGTGCGGATGCGCGGACGCGGCGCGGATCAGCGGTACGCGGCGCATCGCTCGA 94476
QY 457 CCAGTACCATTACCAACAGCTTATTTCCGAGGATGACCAACACACTGCGCAATCCGCT 516

```

Db      94477  CCGGCGTGATCCGGAACGGCTCGTACCATGGCTGTGACGACACATCGGCCGCGCTGT 94536
QY      517  AGACCGGTACCGTGGGAAAGGCGAC 543
      ||| | | | | | | | | | | | | |
Db      94537  AGAACACCGCGACGCTGTCACGCCGC 94563

RESULT 45
LOCUS    ECOW89
DEFINITION E. coli chromosomal region from 89.2 to 92.8 minutes.
ACCESSION U00006
VERSION   U00006.1 GI:409785
KEYWORDS  Escherichia coli (sub_strain MGI655, strain K-12) (library: lambda)
SOURCE   DNA
          Escherichia coli
          Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
          Escherichia.
          1 (bases 1 to 176195)
          Blattner,F.R., Burland,V., Plunkett,G. III, Sofia,H.J. and
          Daniels,D.L.
          Analysis of the Escherichia coli genome. IV. DNA sequence of the
          region from 89.2 to 92.8 minutes
          Nucleic Acids Res. 21 (23), 5408-5417 (1993)
          94083392
          8265357
          2 (bases 1 to 176195)
          Blattner,F.R.
          Direct Submission
          Submitted (03-SEP-1993) 608-263-7459
          On Oct 29, 1993 this sequence version replaced gi:396288.
          This sequence was determined as part of the E. coli Genome Project
          (Frederick R. Blattner, director) at the University of
          Wisconsin-Madison. Supported by award HG0301 from the NIH Human
          Genome Project. The entire sequence was independently determined
          from E. coli MGI655; overlaps and conflicts with other sequence
          determinations are annotated. The start of this entry overlaps the
          end of the entry ECOW87 (L19201) by 93 bp. NOTE: An update was
          submitted on 28-OCT-1993, reflecting a correction to the heme
          sequence.
FEATURES             Location/Qualifiers
     source           1..176195
                     /organism="Escherichia coli"
                     /strain="K-12"
                     /sub_strain="MGI655"
                     /db_xref="taxon:562"
     misc_feature     /note="This sequence comprises the following lambda
                     clones: EC18-126, EC14-54, EC21-52, EC17-24, EC17-203,
                     EC19-105, EC18-208, EC18-110, EC17-200, EC18-33, EC30-262,
                     EC30-41, EC30K637-2a, EC30K637-5, EC27-975, EC22-27;
                     M13mp19 or Janus vectors were used for subcloning"
                     <1..1737
                     /note="corresponds to M21516; ECOKATGA(1225..2805)"
     gene             1..1257
                     /gene="katG"
     CDS              <1..1257
                     /EC_number="1.11.1.6"
                     /standard_name="catalase HPI; catalase-peroxidase; catalase"
                     /note="CG Site No. 14983"
                     /codon_start=1
                     /transl_table=11
                     /product="catalase hydroperoxidase I"
                     /protein_id="AAC43048.1"
                     /db_xref="GI:396289"
                     /translation="ATTSGLVVTOTPTQWSNYFFENLFKVEVWQTRSPAGAIQFEA
                     VDPEIIPDPDSKRRKTMVTDLTAFDPEFKISRRFLNDPAFNAFARXWFK
                     LTRHMGSPRTIGPVPKEDLLWQDLPOPIYNTPEQDIDILKFAIDSGSVSELY
                     SVWASAFRRGDKRGGAARLARLQPDMDVNAARALPVLKIKQESKASLA
                     DITLVAGPKVKAASAGLISHVPPAPGRVMDARQDQDIEMFLEIPEADGPRNRYA
                     RLDSVTESLLIDKAQOLITAPENALVGMRVIGANFDGSKNGVYTDVRGVLSNDF
                     FYNLLDMRYEKATDESKELFEGDRDETGEVKFTASRADLVFGNSVLRVAEVIASS

```

```

misc_feature     DAHEKTFVKDFVAAWVKVMNLDREFDL"
<1..161
/gene="katG"
/note="corresponds to lambda clone EC18-126"
misc_feature     1..93
/gene="katG"
/note="93 bp overlap with end of L19201 (ECOW87)"
misc_feature     156..15760
/note="corresponds to lambda clone EC14-54"
misc_difference   938
/gene="katG"
/note="G in M21516; C here"
1266..1305
/note="putative"
1317..2255
/note="similar to Desulfurolobus ambivalens hypoth. 28.3
kDa protein in sor 3' region"
/codon_start=1
/transl_table=11
/label_ORF_0312
/protein_id="AAC43049.1"
/db_xref="GI:409786"
/translation="MGERKLHTGSLMSAAGSNPLAISGLVLTLLIWSYWIEMKQVT
SYTGADFTALRCIFGALVLFIVLLRGRGMRPTPKYTLIAILLQTCGNYGLAOWAL
VSGCAGKVALISTYMFVWVIFALFLGERLRGQIFAILIAAFGLFVLQPDLPFS
SMKSAMAILSLSGVSGASAIKARLYARHPRVDLLSLTSQWMLYALVMSVALLVPO
REIDQPTVFMALAYSAITALAWLSLWLFVLNKLPAISLASLSTLAVPVGCVLFSWML
LGNPGAVGSGIVLIVLALALVSRKKEAVSVKRI"
1429..1432
misc_difference   /note="TTT in M21516; TTTT here"
1720
misc_difference   /note="C in M21516; G here"
complement(2282..2899)
CDS              /codon_start=1
                 /transl_table=11
                 /label_ORF_f205
                 /protein_id="AAC43050.1"
                 /db_xref="GI:396291"
                 /translation="MKASLALLSLTAFTHSLKSPAVPTVVQIQANTLAIADGAR
                 QQTGSLFDPAYVQITPGDVPQBGVCSVDVIRALRSQKVDLQKLVHEDMAKNFA
                 EYQKWLKRPDSNIDHRRVPLETWFSRHKDTRPTSKNPSDYQAGDISWRDLNGLA
                 HGVVSDGFARDGTPLVIHNIAGAGAEEDVFNWRVGHYRFVK"
                 complement(2886..2915)
                 /note="promoter matrix score of 42; putative; within ORF
                 f205, which would suggest alternate start codon"
                 2912..3053
                 /note="predicted bend of 81 degrees"
                 3031..3064
                 /standard_name="REP; repetitive extragenic palindromic
                 element"
                 /note="contains 1 REP sequence"
                 3090..3219
                 /note="predicted bend of 75 degrees"
                 complement(3174..4316)
                 /gene="gldA"
                 /EC_number="3.174..4316"
                 /gene="gldA"
                 /note="similar to Bacillus stearothermophilus glycerol
                 dehydrogenase"
                 /codon_start=1
                 /transl_table=11
                 /label_ORF_f380
                 /product="glycerol dehydrogenase"
                 /protein_id="AAC43051.1"
                 /db_xref="GI:396292"
                 /translation="MPLHLLISKAIMDRIIOSPGYIOGADVINELGKPLKXER
                 WLWVGKDFVLGAQSVFSEKDFKAGLVEIAPFGCCSQNEIDRLRGIAETACQAGIL
                 GUGGKTLDTAKAHLFMGVPVAIPTASTDAPCSALSIVITDDEGFDYRLLPNP
                 NMVITDKIVAGAPARLLAGIDALATWFARACRSAGTAMAGCKTQAALALAEI
                 CYNLTLEEGEKAMAAEQHVVTIPALERTYIANTYLSGVGFSGGGLAAAHVHGLTAI
                 PDAAHYHGEKAVAGFTLTOLVLENAPVEETVAALSHVGLPITLALDIDKEDVPAK
                 MRLVAEAKCAEGETIHNMPGATPDQVYALLVADQYQGRFLQEW"
4022..4156
misc_structure

```



```

* 14175 15258: contig of 1084 bp in length
* 15259 15358: gap of unknown length
* 15359 16585: contig of 1227 bp in length
* 16586 16861: gap of unknown length
* 16861 18061: contig of 1376 bp in length
* 18062 18161: gap of unknown length
* 18162 19504: contig of 1343 bp in length
* 19505 19604: gap of unknown length
* 19605 21161: contig of 1557 bp in length
* 21162 22375: gap of unknown length
* 22376 22475: contig of 1114 bp in length
* 22476 23942: gap of unknown length
* 23943 24042: gap of unknown length
* 24043 25894: contig of 1852 bp in length
* 25895 25994: gap of unknown length
* 25995 27573: contig of 1579 bp in length
* 27574 27673: gap of unknown length
* 27674 29345: contig of 1672 bp in length
* 29346 29445: gap of unknown length
* 29446 32419: contig of 2974 bp in length
* 32420 32519: gap of unknown length
* 32520 34592: contig of 2073 bp in length
* 34593 34692: gap of unknown length
* 34693 38815: contig of 4123 bp in length
* 38816 38915: gap of unknown length
* 38916 44847: contig of 5932 bp in length
* 44848 44947: gap of unknown length
* 44948 50756: contig of 5809 bp in length
* 50757 50856: gap of unknown length
* 50857 58959: contig of 8113 bp in length
* 58960 59058: gap of unknown length
* 59059 72750: contig of 13681 bp in length
* 72751 72850: gap of unknown length
* 72851 86053: contig of 13203 bp in length
* 86054 86153: gap of unknown length
* 86154 102699: contig of 16546 bp in length
* 102700 102799: gap of unknown length
* 102800 122169: contig of 19370 bp in length
* 122170 122269: gap of unknown length
* 122270 142795: contig of 20526 bp in length
* 142796 142895: gap of unknown length
* 142896 164429: contig of 21534 bp in length
* 164430 164529: gap of unknown length
* 164530 196146: contig of 31617 bp in length.

```

FEATURES

source

```

1. .196146
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="CT7-405C9"
BASE COUNT 52111 a 44572 c 44684 g 51449 t 3330 others
ORIGIN

```

```

Query Match 3.8%; Score 45.4; DB 2; Length 196146;
Best Local Similarity 63.1%; Pred. NO. 0.028;
Matches 70; Conservative 0; Mismatches 41; Indels 0; Gaps 0;
QY 11 TGGAAAAAGCCAAATAAAAAATTCGCCATCCAGCGCTCCAGCTGAAAGTAGGCC 70
DB 3742 TAGAGAAAAACCAAGATAGACAGTGCACACCCAGCGCTTCACGCTAAAGAGGTT 3683
QY 71 TGTCTTCGCGGTATTAATGATGACCGTCCCGGTATTTAAACAATGT 121
DB 3682 TAGCCTGGACTTCTGTGGAATGCATAGTAAGAACCTGTCTTGAAAAAATAT 3632

```

RESULT 47

AC078885/c

LOCUS

```

DEFINITION Mus musculus chromosome 6 clone RP23-117123 strain C57BL6/J,

```

```

WORKING DRAFT SEQUENCE, 44 unordered pieces.

```

AC078885

VERSION

```

AC078885.8 GI:18376842

```

KEYWORDS

```

HTG; HTGS_PHASE1; HTGS_DRAFT.

```

SOURCE

```

Mus musculus

```

ORGANISM

```

Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

```

Eukaryota; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

```

```

1 (bases 1 to 254993)

```

```

Montgomery,K.T., Grills,G., Han,J., Lee,E., Long,J., Pomerantz,R.,

```

```

Ioshikhes,I.P., Shim,C., Decker,J., Thomas,E., Perera,A.,

```

```

Gordon,M., Goltz,J.S. and Kucherlapati,R.

```

```

High Throughput Mouse Sequencing

```

```

Unpublished

```

```

2 (bases 1 to 254993)

```

```

Montgomery,K.T., Grills,G., Han,J., Lee,E., Long,J., Pomerantz,R.,

```

```

Ioshikhes,I.P., Shim,C., Decker,J., Thomas,E., Perera,A.,

```

```

Gordon,M., Goltz,J.S. and Kucherlapati,R.

```

```

Direct Submission

```

```

Submitted (08-AUG-2000) Department of Molecular Genetics, Albert

```

```

Einstein College of Medicine Genome Center, 1300 Morris Park Ave.,

```

```

Bronx, NY 10461, USA

```

```

On Jan 26, 2002 this sequence version replaced gi:14488282.

```

```

-----Genome Center

```

```

Center: Harvard Partners Genome Center

```

```

Center Code: HPGC

```

```

Web site: http://www.hpcg.org/Sequence/mouse.html

```

```

Contact: hpgc@mcg.harvard.edu

```

```

-----Summary Statistics

```

```

Center project name: AAF

```

```

Sequencing vector: pUC18; L08752

```

```

Chemistry: Dye-terminator; Big Dye; 100%

```

```

*Consensus quality: 247096 at least Q20

```

```

*Consensus quality: 242518 at least Q30

```

```

*Consensus quality: 235589 at least Q40

```

```

Estimated insert size: agarose-FP - N/A

```

```

**Estimated insert size: 254133 - sum-of-contigs

```

```

Quality coverage: agarose-FP - N/A

```

```

Quality coverage: 6 x in Q20 bases; sum-of-contigs estimation

```

```

-----

```

```

* NOTE: This is a 'working draft' sequence. It currently

```

```

* consists of 44 contigs. The true order of the pieces

```

```

* is not known and their order in this sequence record is

```

```

* arbitrary. Gaps between the contigs are represented as

```

```

* runs of N, but the exact sizes of the gaps are unknown.

```

```

* This record will be updated with the finished sequence

```

```

* as soon as it is available and the accession number will

```

```

* be preserved.

```

```

* 1 27415: contig of 27415 bp in length

```

```

* 27416 27435: gap of unknown length

```

```

* 27436 57817: contig of 30382 bp in length

```

```

* 57818 57837: gap of unknown length

```

```

* 57838 81637: contig of 23800 bp in length

```

```

* 81638 81657: gap of unknown length

```

```

* 81658 102155: contig of 20498 bp in length

```

```

* 102156 102175: gap of unknown length

```

```

* 102176 121200: contig of 19025 bp in length

```

```

* 121201 121221: gap of unknown length

```

```

* 121222 135073: contig of 13853 bp in length

```

```

* 135074 135093: gap of unknown length

```

```

* 135094 156689: contig of 21596 bp in length

```

```

* 156690 156709: gap of unknown length

```

```

* 156710 166185: contig of 9476 bp in length

```

```

* 166186 166205: gap of unknown length

```

```

* 166206 178798: contig of 12593 bp in length

```

```

* 178799 178818: gap of unknown length

```

```

* 178819 187247: contig of 8429 bp in length

```

```

* 187248 187267: gap of unknown length

```

```

* 187268 192452: contig of 5225 bp in length

```

```

* 192453 192493: gap of unknown length

```

```

* 192513 200084: contig of 7572 bp in length

```

```

* 200085 200104: gap of unknown length

```

```

* 200105 207054: contig of 6950 bp in length

```

```

* 207055 207074: gap of unknown length

```

```

* 207075 212944: contig of 5870 bp in length

```


Db 17616 MTVMCCWRRRRSYNYRSMGAMGMRKSSWSGMRMGSSASSRRCKSASRSSWCSRRMKGM 17617

QY 398 ATCAGGATCCGCCCGCTCTCACTGGCGATAACGGGCACCGGAGACTGACCTTCAGC 457

Db 17676 GSCWSSKMGGSRSRSASSCKGSGRMRRSKRSKSSKRYKRGRKKRSMTKSGSGKSKC 17733

QY 458 CAGTACCATAACAAAGCTTCATTTTCGGAAGGCATGACCCACACATCGCAATCCGGTA 517

Db 17736 WKRSGMTSSCYIYSASSCHMMSSKSCMCCMMKRRCACCYSSMSCTSMYRCWGMK 17799

QY 518 GACCGGTAACTGGGGAAGGCGACCTGCCATTAACACATCTCCGCTCATTCACAGGTG 577

Db 17796 SYTCGCTTCKKCTGYKRSRTWYWMYSWTSKMTMRAACAMMYCTMSTRMMCMMSMCARK 17855

QY 578 TTCTGTCTGCTGACGAGACGCTGCTTCGTATTCTTCAGCGCGCGCCACACAGGAGC 635

Db 17856 GWCMTWKAGATCAGTAGATGAGATAGAAACAGGTCCTGTCGCCACCCAGATC 17913

RESULT 49

AF498405/c

LOCUS

DEFINITION Pseudomonas aeruginosa serotype 014 putative O-antigen biosynthesis gene cluster, partial sequence.

ACCESSION AF498405

VERSION AC104724

KEYWORDS

SOURCE AF498405.1 GI:20559850

ORGANISM Pseudomonas aeruginosa.

Pseudomonas aeruginosa

Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae; Pseudomonas.

REFERENCE 1 (bases 1 to 13413)

AUTHORS Raymond,C.K., Sims,E.H., Kas,A., Spencer,D.H., Kutayavin,T.V., Ivey,R.G., Zhou,Y., Kaul,R., Clendenning,J.B. and Olson,M.V.

TITLE Genetic variation at the O-antigen biosynthetic locus in Pseudomonas aeruginosa

JOURNAL J. Bacteriol. 184 (13), 3614-3622 (2002)

MEDLINE 22053227

PUBMED 12057956

REFERENCE 2 (bases 1 to 13413)

AUTHORS Raymond,C.K., Sims,E.H., Kas,A., Spencer,D.H., Kutayavin,T.V., Ivey,R.G., Zhou,Y., Kaul,R., Clendenning,J.B. and Olson,M.V.

TITLE Direct Submission

JOURNAL Submitted (04-APR-2002) Genome Center, University of Washington, Box 352145, Seattle, WA 98105-2145, USA

COMMENT On May 13, 2002 this sequence version replaced gi:17975253.

FEATURES

source

1. .13413

/organism="Pseudomonas aeruginosa"

/serotype="014"

/db_xref="taxon:287"

<1..104

/note="ORF_1; C-terminal coding region of rpsA"

/codon_start=3

/transl_table=11

/product="RpsA"

/protein_id="AA027622.1"

/db_xref="GI:20559851"

/translation="ENDAMKRLKQVESAGPTTIGDLIRAQMNQ"

241..525

/note="ORF_2; himD/ihfB"

/codon_start=1

/transl_table=11

/product="HimD"

/protein_id="AA027623.1"

/db_xref="GI:20559852"

/translation="MTKSELIERIVTHQGLSARDVELAKTLMLEQMSOALATGDRIEIRGFGSFLHYRAPVRPNPKTGESVRLDGKFVHPFKGKELDRVNEP"

597..815

/note="ORF_3; hypothetical orfA"

/codon_start=1

/transl_table=11

/protein_id="AA027624.1"

QDS

```
/db_xref="GI:20559853"
/translation="MWALYVILVLENSQSVNLSFLGWRTPEGPLSLFVSLAFVLGGL
GVVLSGHWLRLIGSVKGLHLPVREAK"
1167..2216
/notes="ORF_4; wzz; similar to chain length determinant
protein"
/codon_start=1
/transl_table=11
/product="wzz"
/protein_id="AAM27625.1"
/db_xref="GI:20559854"
/translation="MTETLFFQVAKNEEDLIGLRGLWAERKWLIACTILTSVLG
TYALSPVYEAARTAVPPSWSTAGNLTAGNGLTTSVNDVYVSVFRLNLAEETR
NRFYVYLSLSENERGSKDRLYNLSEMEIORPKDERDERTVYIEKHDPGQAA
DWGRYVYLKLAESGEMQDAQOREPEYKRNTOQDIDILRDAKARRDRINLKEA
YQIAESLKLNEPPLIGQDQQLSSIMEGALATMGTGKALRAIRALEERTSDPPIP
ALRLQEQAMYSTLKLQVKNVAVFRODGVAEVPPDPVRPKGILLVIGALSGLVWG
V
MIAFIVMAVRRRAK"
2303..3568
/notes="ORF_5; similar to UDP-glucose/GDP-mannose
dehydrogenase"
/codon_start=1
/transl_table=11
/protein_id="AAM27626.1"
/db_xref="GI:20559855"
/translation="MDIAKAAVGLGYVGLPLAVEFGKGRSVTGFDINHRSIRDLQGY
DSTLEVEKELESVHLKFTSTLSELOECNFFIVTPTIDEHKQPDLTPLVKAESI
AKVLKNDIVYESTVYGATEEVCVPLERESGLRNSDFYGVYSIBTORJDNIALI
VSTIKVSTPEVAEIVDSLYREIITAGTHKAESIKVAEAAKVIEMTORJDNIALI
NELAIFNKLSEIDTESVLOAAGTWNLPFPLPGVGHGCIQVDPYILTHKAQSIGHP
EILLAGRLNDGMGYVYVQLVKAHLKRIHVDGARVLVGLTFKENCPLRNTKVDY
IVRELAENIADVDVWPVSVVEEAQHEYGTLTPIISAPVEGNYDAVVLVAHAHEFKE
DKIRAFGLSVIYDLKFLDKSDSLRL"
3578..4603
/notes="ORF_6; similar to NAD dependent
epimerase/dehydratase family"
/codon_start=1
/transl_table=11
/protein_id="AAM27627.1"
/db_xref="GI:20559856"
/translation="MNAKAKSLRLETTSSWLVTVGAGFTGSLNLELLKFNOKVVG
LDDSTGHYNLDEVRLGVGERWNEFKLOGDIRIDGDCRAVEGVYDVLHQALGS
VPSIADPITNSVNSGLNLVNAKREAGVKSLTYAASSTYGDHPLKVERVIGK
PLSPYATKYVVELYADVFSYGNLGLIRYFNVGRQDPNGAYIAVIFKWLISAMI
KGEPIYINGDSTSDCFINNVKANLLAFCETEAKNQVINYVAGGRTTIGQLDF
SLKVELSRNGINYEQGIYREFRAGDVHRSQAEIKKAQAYLGYNPEYDIQAGIAEMP
WYISFVK"
4665..5954
/notes="ORF_7; similar to Polysaccharide biosynthesis
protein; potential multiple membrane spanning domains."
/codon_start=1
/transl_table=11
/protein_id="AAM27628.1"
/db_xref="GI:20559857"
/translation="MNPKNIFNLGILGLPIIALVPHLIATVGVRFGLFALSNG
LIGAGILDLGGRVATOKLSIRGSDSEMSYTIMTAIRITAVGGLGSLIIFG
LIGGFEFFSHKDVSSLEFVLLAFALPQASATYAGVNEAYLNPKGNLRIPL
GATNFGGFLVSVFYAKDLHYLTIVLSRALAFFIFRLAHGVLKLETVERSCYDR
QAVELFRFGCVTVSSIIINFLVQSDREFIGVLLSAAAVSYVIPYEITQSMILVGA
VSTVAFPSISNLIRTSFAEALCEFNKLLIRVLLIMGGMCLAFLLPFIKLWYGDYI
GDSISVGRILCIGVFENALGAMFYSELHANVKETAILHSIELPIFILILILIPR
HGIVGNANWSLRTLADITVLAVLSYFSWRVKFANC"
6001..7175
/pseudo
/codon_start=1
/transl_table=11
/7156..8091
/notes="ORF_9; similar to Glycosyl transferase"
/codon_start=1
/transl_table=11
/db_xref="GI:20559858"
/translation="MKSPIIDVWVWNNAGLQLRSCVDSIFKNSFGLIRKVIDNGS
```

```
TDNSCALGERENLEVVLTGENTGATACNIGASKGEAEYILFLNPDALYEETIGSV
VFMNRSNSVGVCGVQLIDIEQGVKSRCAFPPSSARMLMHTMGVDRVPFALGYMT
DWHLSSRPDVHVGAFYFRRHVFKLVGGDEFVYLEDLDLSYISCSGSSFYL
STAOAFHAGGASRQIKARRVLSRLIYIFKNLPVWSAFILLLSSLFEPVIRSM
FSWKLSLPALRETWASVLMYLRWLPKWIFFNATR"
8097..9167
/notes="ORF_10; similar to Glycosyl transferases group 1"
/codon_start=1
/transl_table=11
/protein_id="AAM27630.1"
/db_xref="GI:20559859"
/translation="MKILLISKYPLRGASSRLTQYIDYLESKGRVDVRSFFDDGY
LSGLYSSRRSCNWLKSYISRFVFLVYRDLINWKEIFPFPPSVAERFLREFGK
PYVDYDAIFHTNLDLSRNLVFKLKKIDSVMHKSSCIAGNDILAEARVAGSSR
VBIITFDILKLYREYRDESCFKTKTIGWISPTQKLLDIAPELVKLCVQASMLL
VAGTPEISYSGVYVDVNAKNEDEYIRQMTIGIMPLDGPWEKGCCKYKLIQYM
ACGVPVASPGVNIIEVNNESGLLAEGSRGWRESLQKLLQSSSELRECCGRNRSSV
EKIYSLQVSRBELHQTFFIEVLS"
9610..10377
/notes="ORF_11; similar to Glycosyl transferases group 1"
/codon_start=1
/transl_table=11
/protein_id="AAM27631.1"
/db_xref="GI:20559860"
/translation="MAKGSKASVFRSGIAWLYRRALGKKRLVIFQNDDDRDLIGLG
AITEFKSVLRSGVDLSLCHPVPSPGAPVYTLAARLLRQGVYFVEARVAILRQRC
VSAHQVLGDLDPGNPSSIEEFDLEWRSEGVICIGYRQDIASFASHIYVLSYR
EGLPKVLEPAAACGVVYVTDVPGCRDAIEADRTGLLPVRDAVALADAIOKLVESE
LKKMGCAAGRTLAERDFAIESIVQOHLDIYRAGLSGA"
10395..11336
/notes="ORF_12; similar to NAD dependent
epimerase/dehydratase family"
/codon_start=1
/transl_table=11
/protein_id="AAM27632.1"
/db_xref="GI:20559861"
/translation="MTGATGFVGAALVESLCSSEKYRWAGYRRESSPWPEGVTLPLF
GELGSSVAMETESVIDTVHCAARVHMSETVSDPLAEFRKANVLGTLDLARQATSLG
VRFYILSIKYNVEGEGEPQYVADSLPNVDVYGVSKLAEQALGLAAATGLLEV
IIPVLVYGVGVKANYQTMRLKRGVPLPLCAIHNRSLVALNDLTOLIANCIDHPA
ANQVFLSDGEDLSTDLRRNGNALRTPARLPVPVPSWIGLVAKTLNKQAFARLC
GSLQVDIERNRMLGWTTPGVDELEKARYFLDHQ"
11333..12373
/notes="ORF_13; similar to Glycosyl transferase group 4;
potential multiple membrane spanning domains."
/codon_start=1
/transl_table=11
/protein_id="AAM27633.1"
/db_xref="GI:20559862"
/translation="MMSLWLLPMAVAGLSLLTAGLRRYAIARSLMDVPNARSSHQVP
TPRGGVAVIVLSFLAVLAAFLGMDISLAIGLGVGTALGLDHRGHIAARWR
LIGHFAGACWLYLGLPALVFFGMIDLVGWHIGVAVYLVWMLNLPNMDIGDI
ASVEAVCVICAVILLYVVSVCSDASOGIWLISALLASAVTGTFLWFPAPFLMGDA
GSGFLIILAGLSLQAAWSPQLFWGVLILLGVYVDAILTLRLLRLLRGDKYEAHS
HAYQASRYTGRHLPVTLVAGCINLVLLPPLALFVAGSRIDGMALLIGTLPLVAL
RFRAGTLEPRVA"
12571..>13413
/notes="ORF_14; similar to N terminal coding region of
Polysaccharide biosynthesis protein wpbm"
/codon_start=1
/transl_table=11
/protein_id="AAM27634.1"
/db_xref="GI:20559863"
/translation="MLDNLRVLLGLPRKRIQVATDIALVWLSLWLAFLVRLGTE
DMISPGGHAWLFAPLVAIPLFIRFGMYRAVMRYLGNDAIJAKAVTISALVLSL
LVWYIRSPVAVVPRSLVFNFWLWSMLLIGLRAMRQYMDYSAQVSPFLNRQDG
LPRVAYIYGAGAGNQLVALRLGRAMPVAFIDDDKOIANRYTAGLRYTAKHQRMI
DETGAQEVLLAIPSATRRARREITLESLEPFLHVRSPMGFMDLASGRVKVDDIQEVDI
ADLIG"
```

BASE COUNT 2835 a 2899 c 3737 g 3942 t
ORIGIN

3.8%; Score 44.4; DB 1; Length 13413;

Query Match

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 17, 2003, 15:59:59 ; Search time 2683.55 Seconds
(without alignments)
16148.066 Million cell updates/sec

Title: US-09-674-277-1
Perfect score: 1489
Sequence: 1 ctgcagtcgagatgaagaag.....ctggagagagcctggtgcac 1489

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 150 summaries

Database :

- 1: gb.ba.*
- 2: gb.htg.*
- 3: gb.in.*
- 4: gb.om.*
- 5: gb.ov.*
- 6: gb.pat.*
- 7: gb.ph.*
- 8: gb.pl.*
- 9: gb.pr.*
- 10: gb.ro.*
- 11: gb.sts.*
- 12: gb.sv.*
- 13: gb.un.*
- 14: gb.vi.*
- 15: em.ba.*
- 16: em.fun.*
- 17: em.hum.*
- 18: em.in.*
- 19: em.mu.*
- 20: em.om.*
- 21: em.or.*
- 22: em.ov.*
- 23: em.pat.*
- 24: em.ph.*
- 25: em.pl.*
- 26: em.ro.*
- 27: em.sts.*
- 28: em.un.*
- 29: em.vi.*
- 30: em.htg_hum.*
- 31: em.htg_inv.*
- 32: em.htg_other.*
- 33: em.htg_mus.*
- 34: em.htg_pln.*
- 35: em.htg_rnd.*
- 36: em.htg_mam.*
- 37: em.htg_vrt.*
- 38: em.sy.*
- 39: em.htgo_hum.*
- 40: em.htgo_mus.*
- 41: em.htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description	
1	1489	100.0	1489	6	AX011297	AX011297 Sequence	
2	1476.4	99.2	92077	1	AF074613	AF074613 Escherich	
3	1476.4	99.2	92077	6	AX191727	AX191727 Sequence	
4	1476.4	99.2	92721	1	AB011549	AB011549 Escherich	
5	1476.4	99.2	92721	6	AX191725	AX191725 Sequence	
c	7	536.6	36.0	12268	1	ECHECCTP	X89017 E.coli 0157
	8	536.6	36.0	13632	1	AF135170	AF135170 Yersinia
	9	536.6	36.0	220050	1	AJ414156	AJ414156 Yersinia
10	384.8	25.8	2587	1	AB017595	AB017595 Legionell	
11	384.8	25.8	5249	1	AF276752	AF276752 Legionell	
c	12	368	24.7	1829	1	ECIS91TP	X17114 Escherichia
	13	357.6	24.0	2238	6	AR098264	AR098264 Sequence
	14	349.4	23.5	11320	1	AE010762	AE010762 Methanosa
15	348.8	23.4	17986	1	AF294823	AF294823 Shigella	
c	16	334.6	22.5	302150	1	AP001510	AP001510 Bacillus
	17	333.6	22.4	3289	1	AF486647	AF486647 Rhizobium
	18	331.8	22.3	10578	1	AE007314	AE007314 Sinorhizo
19	326.2	21.9	4754	1	RLE438039	AJ438039 Rhizobium	
20	323.2	21.7	340857	1	AP003010	AP003010 Mesorhizo	
c	21	314.8	21.1	198050	1	AL646061	AL646061 Ralstonia
	22	309.2	20.8	2904	1	SRFRPS	Y14317 Streptomyce
	23	305.4	20.5	2805	1	ECOKATGA	M21516 E.coli katG
24	305.4	20.5	13840	1	AE000468	AE000468 Escherich	
25	305.4	20.5	96484	1	ECOUW87	L19201 E. coli chr	
26	303.8	20.4	318703	1	AP002567	AP002567 Escherich	
27	303.4	20.4	3018	8	ANI305225	AJ305225 Aspergill	
28	303	20.3	10667	1	AE005625	AE005625 Escherich	
29	302.2	20.3	2885	1	AF027168	AF027168 Caulobact	
30	302.2	20.3	10677	1	AE005967	AE005967 Caulobact	
31	300.6	20.2	2487	1	HMY16851	Y16851 Haloarcula	
32	298.4	20.0	2223	1	AF314111	AF314111 Mycobacte	
33	298.4	20.0	2223	1	AF314115	AF314115 Mycobacte	
34	298.4	20.0	2223	1	AF314117	AF314117 Mycobacte	
35	298.4	20.0	2223	1	AF314120	AF314120 Mycobacte	
36	298.4	20.0	2223	1	AF314121	AF314121 Mycobacte	
37	298.4	20.0	2223	1	MTU40593	U40593 Mycobacteri	
38	298.4	20.0	2223	1	MTU40595	U40595 Mycobacteri	
39	298.4	20.0	2223	1	MTU41305	U41305 Mycobacteri	
40	298.4	20.0	2223	1	MTU41306	U41306 Mycobacteri	
41	298.4	20.0	2223	1	MTU41310	U41310 Mycobacteri	
42	298.4	20.0	2223	1	MTU41313	U41313 Mycobacteri	
43	298.4	20.0	2223	1	MTU41314	U41314 Mycobacteri	
44	298.4	20.0	2225	1	MTU41304	U41304 Mycobacteri	
45	298.4	20.0	2235	6	I61339	I61339 Sequence 1	
46	298.4	20.0	2235	6	I74293	I74293 Sequence 1	
47	298.4	20.0	2297	1	MTU06269	U06269 Mycobacteri	
48	298.4	20.0	2311	1	MTU06258	U06258 Mycobacteri	
49	298.4	20.0	2316	1	MTU06264	U06264 Mycobacteri	
50	298.4	20.0	2326	1	MTU06260	U06260 Mycobacteri	
51	298.4	20.0	2331	6	I61357	I61357 Sequence 20	
c	52	298.4	20.0	16760	1	AE007051	AE007051 Mycobacte
	53	298.4	20.0	44201	1	MTCY180	297193 Mycobacteri
	54	298.2	20.0	2359	1	MTU06262	U06262 Mycobacteri
55	298	20.0	2322	1	MTU06267	U06267 Mycobacteri	
56	297.4	20.0	2326	1	MTU06270	U06270 Mycobacteri	
57	297.2	20.0	3027	1	AF126956	AF126956 Streptomy	
58	297.2	20.0	10230	1	AE000951	AE000951 Archaeogl	
59	297	19.9	2208	1	AB020234	AB020234 Bacillus	
60	297	19.9	2232	1	AB020065	AB020065 Bacillus	
61	297	19.9	2232	1	AB020120	AB020120 Bacillus	
62	297	19.9	2235	1	AB020067	AB020067 Bacillus	
63	297	19.9	2235	1	AB020092	AB020092 Bacillus	
64	297	19.9	2238	1	AB020115	AB020115 Bacillus	
65	297	19.9	2241	1	AB020094	AB020094 Bacillus	

66	297	19.9	2247	1	AB020078	Bacillus	AB020078	139	296.8	19.9	4810	1	MTKATGCP	X68091	M. tuberculosis
67	297	19.9	2250	1	AB020106	Bacillus	AB020106	140	296.4	19.9	2308	1	MTU06266	U06266	Mycobacteri
68	297	19.9	2253	1	AB020116	Bacillus	AB020116	141	296.4	19.9	2324	1	MTU06268	U06268	Mycobacteri
69	297	19.9	2253	1	AB020117	Bacillus	AB020117	142	295.8	19.9	2316	1	MTU06265	U06265	Mycobacteri
70	297	19.9	2256	1	AB020064	Bacillus	AB020064	143	295.6	19.9	14922	1	SCF73	AL121746	Streptomy
71	297	19.9	2256	1	AB020066	Bacillus	AB020066	144	295	19.8	2312	1	MTU06263	U06263	Mycobacteri
72	297	19.9	2256	1	AB020071	Bacillus	AB020071	145	294.8	19.8	2332	1	MTU06272	U06272	Mycobacteri
73	297	19.9	2256	1	AB020072	Bacillus	AB020072	146	293.8	19.7	2208	1	AB020235	AB020235	Bacillus
74	297	19.9	2256	1	AB020073	Bacillus	AB020073	147	293.8	19.7	2238	1	AB020123	AB020123	Bacillus
75	297	19.9	2256	1	AB020074	Bacillus	AB020074	148	293.8	19.7	2238	1	AB020156	AB020156	Bacillus
76	297	19.9	2256	1	AB020075	Bacillus	AB020075	149	293.8	19.7	2244	1	AB020158	AB020158	Bacillus
77	297	19.9	2256	1	AB020076	Bacillus	AB020076	150	293.8	19.7	2256	1	AB020122	AB020122	Bacillus
78	297	19.9	2256	1	AB020077	Bacillus	AB020077								
79	297	19.9	2256	1	AB020079	Bacillus	AB020079								
80	297	19.9	2256	1	AB020080	Bacillus	AB020080								
81	297	19.9	2256	1	AB020081	Bacillus	AB020081								
82	297	19.9	2256	1	AB020082	Bacillus	AB020082								
83	297	19.9	2256	1	AB020083	Bacillus	AB020083								
84	297	19.9	2256	1	AB020084	Bacillus	AB020084								
85	297	19.9	2256	1	AB020085	Bacillus	AB020085								
86	297	19.9	2256	1	AB020086	Bacillus	AB020086								
87	297	19.9	2256	1	AB020087	Bacillus	AB020087								
88	297	19.9	2256	1	AB020088	Bacillus	AB020088								
89	297	19.9	2256	1	AB020090	Bacillus	AB020090								
90	297	19.9	2256	1	AB020091	Bacillus	AB020091								
91	297	19.9	2256	1	AB020093	Bacillus	AB020093								
92	297	19.9	2256	1	AB020096	Bacillus	AB020096								
93	297	19.9	2256	1	AB020097	Bacillus	AB020097								
94	297	19.9	2256	1	AB020098	Bacillus	AB020098								
95	297	19.9	2256	1	AB020099	Bacillus	AB020099								
96	297	19.9	2256	1	AB020101	Bacillus	AB020101								
97	297	19.9	2256	1	AB020102	Bacillus	AB020102								
98	297	19.9	2256	1	AB020104	Bacillus	AB020104								
99	297	19.9	2256	1	AB020105	Bacillus	AB020105								

```

RESULT 2
AF074613          92077 bp      DNA      circular BCT 04-NOV-1998
LOCUS
DEFINITION      Escherichia coli O157:H7 plasmid p0157, complete sequence.
ACCESSION      AF074613
VERSION        AF074613.1 GI:3822114
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS      Burland,V., Shao,Y., Perna,N.T., Plunkett,G., Sofia,H.J. and Blattner,F.R.
TITLE        The complete DNA sequence and analysis of the large virulence plasmid of Escherichia coli O157:H7
JOURNAL      Nucleic Acids Res. 26 (18), 4196-4204 (1998)
MEDLINE
PUBMED      98391744
PUBMED      9722640
REFERENCE
AUTHORS      Burland,V., Shao,Y., Perna,N.T., Plunkett,G. III, Sofia,H.J. and Blattner,F.R.
TITLE        Direct Submission
JOURNAL      Submitted (25-JUN-1998) Genetics, University of Wisconsin, 445 Henry Mall, Madison, WI 53706, USA
FEATURES
source
1. .92077
   location/Qualifiers
   /organism="Escherichia coli O157:H7"
   /strain="EDL933"
   /serotype="O157:H7"
   /db_xref="taxon:83334"
   /lab_host="Escherichia coli C600"
   /plasmid="p0157"
   1. .561
   /gene="f1no"
   /note="L7001"
   1. .561
   /gene="f1no"
   /note="97 pct identical amino acid sequence and equal length to F101_ECOLI SW: P22707"
   /codon_start=1
   /transl_table=11
   /product="fertility inhibition protein (conjugal transfer repressor)"
   /protein_id="AAC70069.1"
   /db_xref="GI:3822115"
   /translation="MAEQKRPVLTLRKTEGETPVRSRKTIINVTTPPKWKVKQKLA EKARAEALAAKQAQQAQALSIYINLPTQDEAVNTLKPMPWGLFDGTPRLACGIRL DVAQRNIPLSHKLRALAKAITSESYLCAMKAGACRYDTEGVVTEHISQEEEA YAAARDKIRQRNRIKAEQLAVLDER"
   653. .681
   /note="predicted sigma 70 promoter; score of 52%"
   698. .949
   /gene="L7002"
   698. .949
   /gene="L7002"
   /note="54 pct identical (1 gap) to 23 residues of an approx. 289 aa protein EXOK_RHIME SW: P33693"
   /codon_start=1
   /transl_table=11
   /product="unknown"
   /protein_id="AAC70070.1"
   /db_xref="GI:3822116"
   /translation="MDETVHGIVRSVTSVPDAGPLFWKSVDAGWKRQKHGDLPLV HPLTGSSLPKELGNTATGAAREGGNEKSLSHYRSDSEG"
   1025. .1052
   /note="predicted sigma 70 promoter; score of 62%"
   1151. .1612
   /gene="L7003"
   1151. .1612
   /gene="L7003"
   /note="98 pct identical and equal length to YF13_ECOLI SW:

```


QY	180	AGGCGGGGGAACCTCCGTCGGGCAGTGAACCGTGCACACTCCGGGCAGTACATGCCGC	239
Db	7016	AGGCGGGGGAACCTCCGTCGGGCAGTGAACCGTGCACACTCCGGGCAGTACATGCCGC	7075
QY	240	CGGCGCTGATACCGGCAAGATGCTCCCAACTCCCGCTCCGTCGACGGCTATTTCAG	299
Db	7076	CGGCGCTGATACCGGCAAGATGCTCCCAACTCCCGCTCCGTCGACGGCTATTTCAG	7135
QY	300	GATACCCCTTCGTATCAACACGTAACAACAGACAGACAGCTTTTGTCTGACATCCA	359
Db	7136	GATACCCCTTCGTATCAACACGTAACAACAGACAGACAGCTTTTGTCTGACATCCA	7195
QY	360	CAAGAAGGGAATATTCAGTCTCGGACAGACTCAACGGCATCGTACGTTGGCTTGGGA	419
Db	7196	CAAGAAGGGAATATTCAGTCTCGGACAGACTCAACGGCATCGTACGTTGGCTTGGGA	7255
QY	420	ACCCCTTAGTATTTTGTCTGTAGTATCTATCCAGCAATAGGTATATCCTGTGCATC	479
Db	7256	ACCCCTTAGTATTTTGTCTGTAGTATCTATCCAGCAATAGGTATATCCTGTGCATC	7315
QY	480	AATAAGTGTGACTTTGTATACACATGCGAATTTCCCTTAATCCGGAGCTATTTCGATG	539
Db	7316	AATAAGTGTGACTTTGTATACACATGCGAATTTCCCTTAATCCGGAGCTATTTCGATG	7375
QY	540	ATAAAAAAACTCTTCTGTGATTTCTCTGGCGTATCCGGGAGCTTTTCTACCGCT	599
Db	7376	ATAAAAAAACTCTTCTGTGATTTCTCTGGCGTATCCGGGAGCTTTTCTACCGCT	7435
QY	600	GTACCCCTGATATAAAGAGACTCAAAATTTCTACTATCCAGAAACACTGGATTAACT	659
Db	7436	GTACCCCTGATATAAAGAGACTCAAAATTTCTACTATCCAGAAACACTGGATTAACT	7495
QY	660	CCTCTGAGATTACACAGCCCTGAATCAAAATCCCTGGGGGCTGATTTGATTATGCCACC	719
Db	7496	CCTCTGAGATTACACAGCCCTGAATCAAAATCCCTGGGGGCTGATTTGATTATGCCACC	7555
QY	720	AGATTTCAACAGCTGGATATGGAGGCTCTGAAAAAGATATCAAAAGATTTCGTGACAACT	779
Db	7556	AGATTTCAACAGCTGGATATGGAGGCTCTGAAAAAGATATCAAAAGATTTCGTGACAACT	7615
QY	780	TCCAGGATTGGTCCCTCGCGGATATGTTGTCATTTATGCTCTTTTATTCGTATGGCT	839
Db	7616	TCCAGGATTGGTCCCTCGCGGATATGTTGTCATTTATGCTCTTTTATTCGTATGGCT	7675
QY	840	TGGCAGCTGCCGGAACATACAGACATATATGTCGGCGGAGCGCCAGTGTGTCTAG	899
Db	7676	TGGCAGCTGCCGGAACATACAGACATATATGTCGGCGGAGCGCCAGTGTGTCTAG	7735
QY	900	CAACGTTTTGAACCGCTGAACAGTGGCCGGATACGTTAATCTGGATAAAGCCGTCGA	959
Db	7736	CAACGTTTTGAACCGCTGAACAGTGGCCGGATACGTTAATCTGGATAAAGCCGTCGA	7795
QY	960	TTGCTGTGGCAGTCAAGAAAAATACGGCTCCAGTATTTCTGGGAGACCTGATGGTC	1019
Db	7796	TTGCTGTGGCAGTCAAGAAAAATACGGCTCCAGTATTTCTGGGAGACCTGATGGTC	7855
QY	1020	CTGACTGGTAAATGTTCCTTGAATCCATGGGATTTAAACCGCTGGGATTTGCTGGCGGA	1079
Db	7856	CTGACTGGTAAATGTTCCTTGAATCCATGGGATTTAAACCGCTGGGATTTGCTGGCGGA	7915
QY	1080	AGAGAAGTACTGGGAGTGGACCTGGTATACGTGGGGGCTTGACAAAGCCTCTTGCA	1139
Db	7916	AGAGAAGTACTGGGAGTGGACCTGGTATACGTGGGGGCTTGACAAAGCCTCTTGCA	7975
QY	1140	GATAACCGGATAAAGCGGAACCTTCAGAACCTTCTCCGCCACCGCAGATGGGACTT	1199
Db	7976	GATAACCGGATAAAGCGGAACCTTCAGAACCTTCTCCGCCACCGCAGATGGGACTT	8035
QY	1200	ATTATGTCAATCTGAAGCCCGGTTGAAAAACCAAGATCCTCTGGCTTCCCGGAAAGAT	1259
Db	8036	ATTATGTCAATCTGAAGCCCGGTTGAAAAACCAAGATCCTCTGGCTTCCCGGAAAGAT	8095

|||||
7256 ACCCTTAGTATTTTGTCTGTAGTATCTATCCAGCAATAGTATATCTCTGTCATC 7315
QY 480 AATAAGTTGACTTTTGTATACACATGCGAATTTCCCTTAATCCGAGCATATCTGATG 539
Db 7316 ANTAAGTTGACTTTTGTATACACATGCGAATTTCCCTTAATCCGAGCATATCTGATG 7375
QY 540 AATAAAAAACTCTCTCTGTTCTGATTTCTCTGCGCTATCGGGAGCTTTTCTACCGCT 599
Db 7376 AATAAAAAACTCTCTCTGTTCTGATTTCTCTGCGCTATCGGGAGCTTTTCTACCGCT 7435
QY 600 GTAGCCGCTGATAAAAAAGACATCAAAATTTCTACTATCCAGAAACATCTGGATTAAC 659
Db 7436 GTAGCCGCTGATAAAAAAGACATCAAAATTTCTACTATCCAGAAACATCTGGATTAAC 7495
QY 660 CCTCTGAGATTACACAGCCCTGGAATCAATCCCTCGGGGGCTGATTTGATTATGCCACC 719
Db 7496 CCTCTGAGATTACACAGCCCTGGAATCAATCCCTCGGGGGCTGATTTGATTATGCCACC 7555
QY 720 AGATTTCAACAGCTGGATATGAGGCTCTGAAAAAGATATCAAAAGATTTGCTGCAACT 779
Db 7556 AGATTTCAACAGCTGGATATGAGGCTCTGAAAAAGATATCAAAAGATTTGCTGCAACT 7615
QY 780 TCCAGGATTTGCTGCGGATATGCTCATTTATGCTCTTCTTATTCGTATGGCT 839
Db 7616 TCCAGGATTTGCTGCGGATATGCTCATTTATGCTCTTCTTATTCGTATGGCT 7675
QY 840 TGGCAGCTGTCGGAACATACAGACATATGATGCGCGGGAGCGCCAGTGTGTGTCAG 899
Db 7676 TGGCAGCTGTCGGAACATACAGACATATGATGCGCGGGAGCGCCAGTGTGTGTCAG 7735
QY 900 CAACGTTTTGAACGCTGAACAGCTGCGCGGATTAACCTTAATCTGATTAAGCCGCTGCA 959
Db 7736 CAACGTTTTGAACGCTGAACAGCTGCGCGGATTAACCTTAATCTGATTAAGCCGCTGCA 7795
QY 960 TTGCTGTGCGCAGTCAAGAAAAATACGCTCCAGTATTTCTGGGAGACCTGATGCTC 1019
Db 7796 TTGCTGTGCGCAGTCAAGAAAAATACGCTCCAGTATTTCTGGGAGACCTGATGCTC 7855
QY 1020 CTGACTGTGTAATTTGTCCTTTGAATCCATGSGGATTTAAACGCTGGGATTTGCTGCGGA 1079
Db 7856 CTGACTGTGTAATTTGTCCTTTGAATCCATGSGGATTTAAACGCTGGGATTTGCTGCGGA 7915
QY 1080 AGAGAAGATGACTGGGAGTGGACCTGCTATGCTGGGGCTGACACAGCTCTTGCA 1139
Db 7916 AGAGAAGATGACTGGGAGTGGACCTGCTATGCTGGGGCTGACACAGCTCTTGCA 7975
QY 1140 GATAACCGGGATAAAACGGGAAACTTCAGAAACCTCTTGGCGCCACGACATGGGACTT 1199
Db 7976 GATAACCGGGATAAAACGGGAAACTTCAGAAACCTCTTGGCGCCACGACATGGGACTT 8035
QY 1200 ATTTATGTCATCTGAAGGCCCGGTTGGAACACAGATCCTCTGGCTTCCGCGAAGAT 1259
Db 8036 ATTTATGTCATCTGAAGGCCCGGTTGGAACACAGATCCTCTGGCTTCCGCGAAGAT 8095
QY 1260 ATCAGGGAAGCTTTTTCAGTATGGCCATGATGATGAGGAGCTGCGGCTGATCGG 1319
Db 8096 ATCAGGGAAGCTTTTTCAGTATGGCCATGATGATGAGGAGCTGCGGCTGATCGG 8155
QY 1320 GGAGGGCATACATTTGTTGTAAGACATGCTGTCAGGCTCTCTGAAAAATGATTTGGCGCA 1379
Db 8156 GGAGGGCATACATTTGTTGTAAGACATGCTGTCAGGCTCTCTGAAAAATGATTTGGCGCA 8215
QY 1380 GGGCTGATGTCACCTGTCGAGAGACGAGGATGCGGATGGAATAAATATGTTGTCAC 1439
Db 8216 GGGCTGATGTCACCTGTCGAGAGACGAGGATGCGGATGGAATAAATATGTTGTCAC 8275
QY 1440 GGAACCGCAATATACCATCAGCTGCTGAGGAGCTGAGGAGCTGCTGTCAC 1489
Db 8276 GGAACCGCAATATACCATCAGCTGCTGAGGAGCTGAGGAGCTGCTGTCAC 8325

RESULT 4

AB011549
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
source
gene
CDS

AB011549 92721 bp DNA circular BCT 27-APR-1999
Escherichia coli plasmid p0157 DNA, complete sequence.
AB011549
AB011549.2 GI:4589740
ToxR-regulated lipoprotein: tagA.
Escherichia coli (strain:O157:H7, sub_strain:RMD 0509952)
plasmid:p0157 DNA.
Escherichia coli
Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
Escherichia.
1 (sites)
Makino,K., Ishii,K., Yasunaga,T., Hattori,M., Yokoyama,K.,
Yutsudo,H.C., Kubota,Y., Yamauchi,Y., Iida,T., Yamamoto,K.,
Honda,T., Han,C.G., Ohtsubo,E., Kasamatsu,M., Hayashi,T., Kuhara,S.
and Shinagawa,H.
Complete nucleotide sequences of 93-kb and 3.3-kb plasmids of an
enterohemorrhagic Escherichia coli O157:H7 derived from Sakai
outbreak
DNA Res. 5 (1), 1-9 (1998)
98290540
2 (bases 1 to 92721)
Makino,K.
Direct Submission
Submitted (24-FEB-1998) Kozo Makino, Research Institute for
Microbial Diseases, Osaka University, Molecular Microbiology,
Yamadaoka, 3-1, Suita, Osaka 562, Japan
(E-mail:makino@biken01.biken.osaka-u.ac.jp, Tel:81-6-879-8318,
Fax:81-6-879-8320)
On Apr 20, 1999 this sequence version replaced gi:3336997.
Location/Qualifiers
1. .92721
/organism="Escherichia coli"
/strain="O157:H7"
/sub_strain="RMD 0509952"
/db_xref="taxon:562"
/plasmid="p0157"
/note="RMD 0509952 is a strain of enterohemorrhagic E.
coli, EHEC O157:H7"
/gene="tagA"
/join(92527..92721,1..2502)
/gene="tagA"
/codon_start=1
/transl_table=11
/product="ToxR-regulated lipoprotein"
/protein_id="BAA31757.3"
/db_xref="GI:4566293"
/translation="MNTKMERWRTPMKLYLSCITLAPLIGVFSATAADNNSAIYF
NTSQPINDLQGSIAAEVFAQSOLPAHPKEDGSDPHLSLRKSLILVPRPVKADKTP
VQVAAEDNNKILGTLTLYPPSSPDITVHLDCVPEGIDFTPHNGTKKILNTVAEYN
KLSDAGSSISHLNNALVEIHTANGRWVRDIYLPQPDLEGRKMRVFSAGYSSTV
FYGDKRVTLSVGTLLFKYNGWFRSELENRIYTAQHISALPARHVIPTGLNLV
IKQGLSGRLNDIKGAPGELLHTIDIGMLTTPDRDFDFAKDKAHEHYFOTIPVSR
MIVNNTYAPLHKEVMTGELATDMDPGNGWHSGTMRQIGKELVSGHIDNANYGLN
STAGLGNHPYVQAALAHNSRGNKANGIQVHGGGGGIWTLDTLGNESHEVGH
NTGLGIVDFGKSVHRSRNNNTNMGDGRKRTIPNFPSTNKSCLNNQCEPF
DGHKFGDAMAGSGPFAANRTMTTPNSSAIIQRFENKAVFDSRSTSGFSKNWADT
QMEPEYHTIDAEQITASVNSLSESKMAELMAEYAVVKVHMNDRVNTNRIYIPTASA
DNRGSTLTINDEAGNYSYLFINGDEKVSQYKKSQVSDGQFWRKRDVVDREARKPE
QFGVPVTVLVYDPEGLTSSYIYPAMYGATGYTSDSONSDNDCLQVDTREGQL
REFLANHRANTVMKFNINPTSOPTQATLVCKNKLIDTKSLTPAPEGLTYTVNGQ
ALPAKNECCIVSVNSGKRYCLPVGQRSGYSLPDWIVQGEVYVDSKAKAKVLLSDWDN
LSYNRIGRFGVGNVPADMKKRWANGQYLDPSKPSMRVYIK"
2589..3464
/gene="etpC"
2589..3464
/gene="etpC"
/codon_start=1
/transl_table=11
/product="Type II secretion pathway related protein"
/protein_id="BAA31758.1"
/db_xref="GI:3336999"

```

/translation="MLFSLFRCDCGLFIKIDIVLKMILTPNRLLCVILLIAGQLVSVI
HPFWLQAASVGLSRVSAPEATAVGTQTEERFVFTLGRASPLSEGRAQETMPSLIS
DDLSSGDLVGLILYSSVAEHSVAIFAHNNQFSLSGVEKVPSTADTAISAFSDHIV
INQGTSLPLRYDNTKKNAYDNNNLTVGDVITQDNFRVESFDIMFSAVTVNNT
LSYRLIPGKSHSLFYNAGLHNDLAVSVNGSELRTQAQIQIMKQLPELKEIKITVE
RGQLYDAFIANGEN"
3675..5432
/gene="etpD"
CDS
3675..5432
/gene="etpD"
/codon_start=1
/transl_table=11
/product="Type II secretion pathway related protein"
/protein_id="BAA31759.1"
/db_xref="GI:3337000"
/translation="MLNEEQYQFFLSVLVDYGFAYVDMHGILKVVRSKDARTSAVP
VASDSPGTGDEVTVTVVPVSNVAARDLAPLRQLQNDNAGASVHVHPESNVLLMTGR
AAVMKRLMEIVERDKVGRNSVATVPLTVASATDVARLVTELTKEVDKTAIPAMWTAK
LVADERTNSVLVSGETISOORLIISIIKOLDROEDVQGNPKVILKYAKAKDLVEVLTG
ISSIENDSKSPSTALAKGYIKSHQETNALILUGADVDINDENVISQDRIKRPQ
VLVEAIIAQDQAGLNLGIQWNRHAGVAQFTSGLPITTVQTRQNEILSDQSNNA
LSMFGIAAGFTQGNWAMLLTALSTSSKNIDILATPSIVTLDNMEATFNQGEVPLSG
SOTTSQDNFENTVERTGVIKLRVKPQINEGSDVLEIEQEVSGVADTAVATVTLGGA
TENTRTVAMLVGNGETVTVVGLLDKSTRGSESKVPLIGDIPVLGHLFRKSEQTAK
RNLMFIRPTIIRERDGFHSAEAKYQSFNQEQVQSRGKETTALTFLNEQLRLSPDQD
DPAFRVKAIAIAFYAQM"
5432..6937
/gene="etpE"
CDS
5432..6937
/gene="etpE"
/codon_start=1
/transl_table=11
/product="Type II secretion pathway related protein"
/protein_id="BAA31760.1"
/db_xref="GI:3337001"
/translation="MSRVVNVSESRLPLPFSEFRTQRIILLREQGNRVFCMEDTPA
SALLEVRRVAEGLPNAVTVVSAEAFKQLVSSYQROSDSEARQMAEIGNEMDFTYVAGE
LRDRELLDANDADPIIRLINLAETAIKEKASDIHETIERHLQVRFKIDGVLRIL
LPHNLSALLISIRKVMARDIDAEKRVQDGRVLRIGGRADVVRVSTPLSNHGERIV
LRLDKNSVSLDLAALGQQNHIDALIRPHGILVLTGPTGSKSTTLAAALSLL
NPRDKNIMVEDPEVELDGIQOTQVNPVDMTFAIRLAILRQDPDVLVGEIRDEGE
TAQIAVOASLTGHLVSLTHTNSAAGALSLODMGIPFPLSTSLIAVAQRLVRLTIC
PRCOPCOVSTELAMDIDIPETTLPQAGCOCQSFYGHGTGTHELLIIDDRITA
IVQGEGLAHTRLAGSRYLTLRGDRQKVLAGETSWEEVVRVTESRLQEE"
6939..8162
/gene="etpF"
CDS
6939..8162
/gene="etpF"
/codon_start=1
/transl_table=11
/product="Type II secretion pathway related protein"
/protein_id="BAA31761.1"
/db_xref="GI:3337002"
/translation="MALPHYQASDIHGRKRSGLILEADSARHAROLLREQALIPVRIDE
KOVHKKHSILIRPRPGSSAEALALLPRLATLVAASLPLEEALDALLROSEKPRQ
RNLLAAYRVKYLEGHSIAAAMGPFCTFERLYCANVAAGETSGRLDVLVLSRLADYTEQ
RQIMRNLQALLIPCVLTAVGVJIAILLTAVVPKVEQFIHMKTLPRLSRVLGMA
AEVSOQWPPWLLAALGAGIAGRMILHQPQSLAFHLLRLRPVVGRIISGLNTARYA
RTLSILNAPLQAAMHISGVDLSNDWARHOLATAAEELVREGVLSLHOLEQTSLPFP
MKRHM1ASGENSELDSMLERAADNQDREFTSQMLALGLFEPFLVVGAGVVLFIVL
ALLOPLLOLNNMNM"
8193..8627
/gene="etpG"
CDS
8193..8627
/gene="etpG"
/codon_start=1
/transl_table=11
/product="Type II secretion pathway related protein"
/protein_id="BAA31762.1"
/db_xref="GI:3337003"
/translation="MRKQHORGFITLLEIMVIVILGVLASLVVPLNMGNKDADROKV
MSDLVALESTLDYRLDNNRYPTTEQGLRALYVSKPTVQPEPRYRQDGIIRRLPQDPW
GGDYQLNFGQISDIDIFSPGPDGVPENTEDDIGNMTLGNAPQ"

```

```

gene
8624..9178
/gene="etpH"
CDS
8624..9178
/gene="etpH"
/codon_start=1
/transl_table=11
/product="Type II secretion pathway related protein"
/protein_id="BAA31763.1"
/db_xref="GI:3337004"
/translation="MSQRGFTILEMMLILLMGTAGLVLMSFPDQAONHQQOORERL
OALDYALDRSQDGLLGMIOVRFDGMEKFKVLQGRGTAESPPTLAEGDILWQGVWQTV
OPRRAMGKGLPDDVRLLEQLYLRGLQWMSHDDGAEPDILLPLPGGVTFPRLLFRQVG
EBAVGLQVDENGLMTLFEGEVSL"
9193..9540
/gene="etpI"
CDS
9193..9540
/gene="etpI"
/codon_start=1
/transl_table=11
/product="Type II secretion pathway related protein"
/protein_id="BAA31764.1"
/db_xref="GI:3337005"
/translation="MTLLEVIVLVFALAGMALMOASTQQAAGIGRMEKVLGWLIA
DNQVQLQLEKTPENGWGEKTIISFAGTIEWYLWNGSPDSDVQPQSRSEVRYRKEET
SALVSRSSVVR"
9537..10136
/gene="etpJ"
CDS
9537..10136
/gene="etpJ"
/codon_start=1
/transl_table=11
/product="Type II secretion pathway related protein"
/protein_id="BAA31765.1"
/db_xref="GI:3337006"
/translation="MSQORVKGFTLEMLLALAVFAALSIASFOVLGGIRAHLSRD
KVQLAEALRGISQMERDLTQMLPHSRGNEGILLAAHPHLLKSDMDGIGIFTRNSWLP
AGMLPPELQWVGVIYRLRQOKLERLSYFHVDPHSGVSPDVVRVMDLGVHAPRLRFFVNGD
WOARMDSTGILPQAEVTVLMDDDFAELPRLFLVSKETA"
10133..11110
/gene="etpK"
CDS
10133..11110
/gene="etpK"
/codon_start=1
/transl_table=11
/product="Type II secretion pathway related protein"
/protein_id="BAA31766.1"
/db_xref="GI:3337007"
/translation="MKLRQGVALLVLLLSLMTVIAAVIAERNRGTFLRTVAQLDQ
LQAKWDGYAETIAKQILQORSQESPRKTHLAQNWASERQFETRGDVRGQIVDAQA
CPNLNAINTVVDLTSIPYAAARIFQOLLINLQVELQAROVTAALDRDDPKVVRG
GAEDVYMGMEPPYLAANQMDQVSELRIGIDARLYKLLPYVCLVTSLSVNVN
TLIDSQAPLILAAALFTKPDLSLPVTELLQRRPRTGWESVAAFLDAPLKDIDTSAAMPV
LAVSNYFVLRHLVRSGEHLFSQOTLMQWREFRFRIIQRYGLTNEVP"
11473..12321
/gene="etpL"

```

```

Query Match          99.2%; Score 1476.4; DB 1; Length 92721;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1488; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 CTGCAGTCGGAGATGAA-AGCACCACTGTGTGTACCCCATCAGCGTGGTCCCGCAGGCC 59
Db 76351 CTGCAGTCGGAGATGAACAGCACCACCTGTGTGTACCCCATCAGCGTGGTCCCGCAGGCC 76410
QY 60 ATGATTTTGTACAGACTCAATGACTACCGGACGACGTGAACCTTCCCGTGTCTTC 119
Db 76411 ATGATTTTGTACAGACTCAATGACTACCGGACGACGTGAACCTTCCCGTGTCTTC 76470
QY 120 AGCCAGTTAAGCCAGCGGTTTCCTGCTGTAATAATGTCGCCAAACCGGGAAGCATCAGA 179
Db 76471 AGCCAGTTAAGCCAGCGGTTTCCTGCTGTAATAATGTCGCCAAACCGGGAAGCATCAGA 76530
QY 180 AGGGCGGGGAAATCCGTCGGGCCAGTGAACCGTGCACACTCCGGGCGAGTACATCCCGC 239

```

```
|||||
76531 AGGCGGGGAACTCCGTCGGCGCAGTGAACCGTGCCACACTCCGGGACGTACATCCCG 76590
QY 240 CGGCGCTGATACCGGCAAGATGGTCCAAACTCCCGCTCCGTCGGAGCGGCTATTTCAG 299
Db 76591 CGGCGCTGATACCGGCAAGATGGTCCAAACTCCCGCTCCGTCGGAGCGGCTATTTCAG 76650
QY 300 GATACCTTCGCTCATCAACAGCTACAAACAGAGACACAGCTTTTGTCTTCTCAGATCCA 359
Db 76651 GATACCTTCGCTCATCAACAGCTACAAACAGAGACACAGCTTTTGTCTTCTCAGATCCA 76710
QY 360 CAAGAAGGGAATATTACGCTCTCGCAGCACTCAACGGCATCGTCAGTTGGCGCTTGA 419
Db 76711 CAAGAAGGGAATATTACGCTCTCGCAGCACTCAACGGCATCGTCAGTTGGCGCTTGA 76770
QY 420 ACCCTTAGTATTTTGTCTGCTAGTATCTATCCAGCAATAGTATATCTTGTGCATC 479
Db 76771 ACCCTTAGTATTTTGTCTGCTAGTATCTATCCAGCAATAGTATATCTTGTGCATC 76830
QY 480 AATAAGTTGACTTTTGTATACACATGCGAATTTCCCTTAATCCGGAGCTATTTCGTATG 539
Db 76831 AATAAGTTGACTTTTGTATACACATGCGAATTTCCCTTAATCCGGAGCTATTTCGTATG 76890
QY 540 ATAAAAAATCTTCTCTGCTGCTGATCTTCTGCGCTATCGGGAGCTTTTCTACCGCT 599
Db 76891 ATAAAAAATCTTCTCTGCTGCTGATCTTCTGCGCTATCGGGAGCTTTTCTACCGCT 76950
QY 600 GTAGCGCTGATAAAAAAGACATCAAAATTTCTACTATCCAGAAACACTGGATTAACT 659
Db 76951 GTAGCGCTGATAAAAAAGACATCAAAATTTCTACTATCCAGAAACACTGGATTAACT 77010
QY 660 COTCTGAGATTACAGCGCTGAATCAATCCCTGGGGGCTGATTTGATTATGCCAACC 719
Db 77011 COTCTGAGATTACAGCGCTGAATCAATCCCTGGGGGCTGATTTGATTATGCCAACC 77070
QY 720 AGATTTCACAGCTGGATATGGAGCTCTGAAAAAGATATCAAGATTTCGTGACAACT 779
Db 77071 AGATTTCACAGCTGGATATGGAGCTCTGAAAAAGATATCAAGATTTCGTGACAACT 77130
QY 780 TCCAGGATTGGTCCCTGCGGATTATGTCATTATGTCCTTTCTTTATTCGTATGGCT 839
Db 77131 TCCAGGATTGGTCCCTGCGGATTATGTCATTATGTCCTTTCTTTATTCGTATGGCT 77190
QY 840 TGGCAGCTGCGGGAACATACAGACATATGATGGCGGGAGCGCCAGTGGTGTGAC 899
Db 77191 TGGCAGCTGCGGGAACATACAGACATATGATGGCGGGAGCGCCAGTGGTGTGAC 77250
QY 900 CAACGTTTGAACCGCTCAACAGCTGGCGGATACGTTAATCTGGATAAAGCCCGTCA 959
Db 77251 CAACGTTTGAACCGCTCAACAGCTGGCGGATACGTTAATCTGGATAAAGCCCGTCA 77310
QY 960 TTGCTGTGCGCAGTCAAGAAAAATACGGCTCCAGTATTTCTGGGAGACCTGATGGTC 1019
Db 77311 TTGCTGTGCGCAGTCAAGAAAAATACGGCTCCAGTATTTCTGGGAGACCTGATGGTC 77370
QY 1020 CTGACTGGTATGTTGGCCCTTGAATCCATGGGATTTAAACGCTGGGATTTGCTGGCGGA 1079
Db 77371 CTGACTGGTATGTTGGCCCTTGAATCCATGGGATTTAAACGCTGGGATTTGCTGGCGGA 77430
QY 1080 AGAGAAGATGACTGGGAGTCCGACCTGCTATCTAGTGGGGCTGACAAACAGCTCTTTGCA 1139
Db 77431 AGAGAAGATGACTGGGAGTCCGACCTGCTATCTAGTGGGGCTGACAAACAGCTCTTTGCA 77490
QY 1140 GATAACCGGGATATAAAGGGAACCTTCAGAAACCTCTTGCGGCACGACGATGGGACTT 1199
Db 77491 GATAACCGGGATATAAAGGGAACCTTCAGAAACCTCTTGCGGCACGACGATGGGACTT 77550
QY 1200 ATTTATGCTCAATCTGAAGGCCCGGCGGAAACAGATCTCTGCGCTTCGCGCAAGAT 1259
Db 77551 ATTTATGCTCAATCTGAAGGCCCGGCGGAAACAGATCTCTGCGCTTCGCGCAAGAT 77610
QY 1260 ATCAGGGAGCTTTTTCAGCTATGGCCATGATGATGAGGAGACTGTGGCCCTCATCGCG 1319
|||||

Db 77611 ATCAGGGAGCTTTTTCAGCTATGGCCATGATGATGAGGAGACTGTGGCCCTCATCGCG 77670
QY 1320 GGAGGGCATACATTTGGTAAAGCATGTCAGCTCTCTGAAAAATGATTATGGCGCA 1379
Db 77671 GGAGGGCATACATTTGGTAAAGCATGTCAGCTCTCTGAAAAATGATTATGGCGCA 77730
QY 1380 GGGCCTGATGTGTGACCTGTGTGAGGAGCAGGACTGGGATGAAAAATGAAATGTTGGTACA 1439
Db 77731 GGGCCTGATGTGTGACCTGTGTGAGGAGCAGGACTGGGATGAAAAATGAAATGTTGGTACA 77790
QY 1440 GGAACGGCAATATATACCATCACAGTGGCTTGGAGGAGCTGGTGCAC 1489
Db 77791 GGAACGGCAATATATACCATCACAGTGGCTTGGAGGAGCTGGTGCAC 77840

RESULT 5
AX191725 92721 bp DNA linear PAT 15-AUG-2001
LOCUS Sequence 7 from Patent WO0149775.
DEFINITION AX191725
ACCESSION AX191725
VERSION AX191725.1 GI:15209894
KEYWORDS
SOURCE Escherichia coli.
ORGANISM Escherichia coli.
Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
Escherichia.
REFERENCE
1 (bases 1 to 92721)
AUTHORS Iversen, P.L.
TITLE Antisense antibacterial cell division composition and method
JOURNAL Patent: WO 0149775-A 7 12-JUL-2001;
Avi Biopharma, Inc. (US)
FEATURES
Location/Qualifiers
source
1..92721
/organism="Escherichia coli"
/db_xref="taxon:562"
BASE COUNT 24870 a 20267 c 23868 g 23716 t
ORIGIN

Query Match 99.2%; Score 1476.4; DB 6; Length 92721;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1488; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 CTGCACTCCGGAGATGAA-AGCACCACTGTGTGTACCCCATCAGCTGGTCCCGCAGGCC 59
Db 76351 CTGCACTCCGGAGATGAAACAGCACCACTGTGTGTACCCCATCAGCTGGTCCCGCAGGCC 76410
QY 60 ATGATTTTGTGCACAGACTCAATGACTACCGGAGCAGCTGAACCTTCGCGTTGTTTCTCC 119
Db 76411 ATGATTTTGTGCACAGACTCAATGACTACCGGAGCAGCTGAACCTTCGCGTTGTTTCTCC 76470
QY 120 AGCCAGTTAAGCAGCGGTTTCCCTGCTGAAAAATGTGCGGAAAAACGGGGAAGCATCAGA 179
Db 76471 AGCCAGTTAAGCAGCGGTTTCCCTGCTGAAAAATGTGCGGAAAAACGGGGAAGCATCAGA 76530
QY 180 AGGCGGGGAACTCCGTCGGCCAGTGAACCTGCCACACTCCGGGCACTGATCATGCCGC 239
Db 76531 AGGCGGGGAACTCCGTCGGCCAGTGAACCTGCCACACTCCGGGCACTGATCATGCCGC 76590
QY 240 CGGCGCTGATACCGGCAAGATGGTGCAGAACTCCCGCTCCGTCGAGCGGCGCTATTTCAG 299
Db 76591 CGGCGCTGATACCGGCAAGATGGTGCAGAACTCCCGCTCCGTCGAGCGGCGCTATTTCAG 76650
QY 300 GATACCTTCGCTCATCAACAGCTACAAACAGAGACAGCTTTTGTCTTCTGACATCCA 359
Db 76651 GATACCTTCGCTCATCAACAGCTACAAACAGAGACAGCTTTTGTCTTCTGACATCCA 76710
QY 360 CAAGAAGGGAATATTACGCTCTCGCAGCACTCAACGGCATCGTCAGTTGGCGCTTGA 419
Db 76711 CAAGAAGGGAATATTACGCTCTCGCAGCACTCAACGGCATCGTCAGTTGGCGCTTGA 76770
QY 420 ACCCTTAGTATTTTGTCTGCTAGTATCTATCCAGCAATAGTATATCTTGTGCATC 479
Db 76771 ACCCTTAGTATTTTGTCTGCTAGTATCTATCCAGCAATAGTATATCTTGTGCATC 76830
```

QY	480	ATAAAGTTGACTTTTGTATACAAATCGAATTTCCCTTAATCCGGAGCTATTCGTATG	539
Db	76831	ATAAAGTTGACTTTTGTATACAAATCGAATTTCCCTTAATCCGGAGCTATTCGTATG	76890
QY	540	ATAAATAAATCTTCCCTGTTCTGATCTTCTGCGCTATCCGGAGCTTTTCTACCGCT	599
Db	76891	ATAAATAAATCTTCCCTGTTCTGATCTTCTGCGCTATCCGGAGCTTTTCTACCGCT	76950
QY	600	GTAGCGCTGTATAAAGAGACTCAAAATTTCTACTATCCAGAAACACTGGATTAAC	659
Db	76951	GTAGCGCTGTATAAAGAGACTCAAAATTTCTACTATCCAGAAACACTGGATTAAC	77010
QY	660	CTCTGAGATTACAGCCCTGAATCAATCCCTGGGGGCTGATTTGATTATGCCAC	719
Db	77011	CTCTGAGATTACAGCCCTGAATCAATCCCTGGGGGCTGATTTGATTATGCCAC	77070
QY	720	AGATTTCAACAGCTGGATGAGGCTCTGAAAAAGATATCAAAAGATTGCTGACAAC	779
Db	77071	AGATTTCAACAGCTGGATGAGGCTCTGAAAAAGATATCAAAAGATTGCTGACAAC	77130
QY	780	TCCAGGATGTCGCCCTCGGATGATGTCATTTATGTCCTTTCTTATTCGTATGGCT	839
Db	77131	TCCAGGATGTCGCCCTCGGATGATGTCATTTATGTCCTTTCTTATTCGTATGGCT	77190
QY	840	TGCGAGGTGCGGCAACATACAGGACATATGATGCGCGGGAGCGCCAGTGGTCTAG	899
Db	77191	TGCGAGGTGCGGCAACATACAGGACATATGATGCGCGGGAGCGCCAGTGGTCTAG	77250
QY	900	CAAGCTTTTGAACCGCTGAACAGCTGGCGGATACGTTAATCTGATAAAGCCGTCGA	959
Db	77251	CAAGCTTTTGAACCGCTGAACAGCTGGCGGATACGTTAATCTGATAAAGCCGTCGA	77310
QY	960	TTGCTGTGCCAGTCAAGAAAAATACGCTCCAGTATTTCTTGGGGAGACCTGATGTC	1019
Db	77311	TTGCTGTGCCAGTCAAGAAAAATACGCTCCAGTATTTCTTGGGGAGACCTGATGTC	77370
QY	1020	CTGACTGTGAATTTGTCCTTGAATCCATGGATTTAAACGCTGGATTTCTGGCGGA	1079
Db	77371	CTGACTGTGAATTTGTCCTTGAATCCATGGATTTAAACGCTGGATTTCTGGCGGA	77430
QY	1080	AGAGAAGATGACTGGAGTCGAGCTGGTATGTTACTGGGGCTTGACAAAGCCTCTTGA	1139
Db	77431	AGAGAAGATGACTGGAGTCGAGCTGGTATGTTACTGGGGCTTGACAAAGCCTCTTGA	77490
QY	1140	GATAACGGGATATAAAGCGGAACCTTCAGAACCTCTTGGCGCCACGAGATGGACT	1199
Db	77491	GATAACGGGATATAAAGCGGAACCTTCAGAACCTCTTGGCGCCACGAGATGGACT	77550
QY	1200	ATTATGTCATCTGAAGCCCGGTGGAACACAGATCCTTCTGCTTCCCGCAAGAT	1259
Db	77551	ATTATGTCATCTGAAGCCCGGTGGAACACAGATCCTTCTGCTTCCCGCAAGAT	77610
QY	1260	ATCAGGAAGCTTTTTCAGTATGGCCATGGATGATGAGGACTTGGCCCTGATCGG	1319
Db	77611	ATCAGGAAGCTTTTTCAGTATGGCCATGGATGATGAGGACTTGGCCCTGATCGG	77670
QY	1320	GGAGGCATACATTTGGTAAGCACATGTCAGGCTCTCCGAAAAATGATTTGGCGCA	1379
Db	77671	GGAGGCATACATTTGGTAAGCACATGTCAGGCTCTCCGAAAAATGATTTGGCGCA	77730
QY	1380	GGGCTGTATGTCACCTGTGGAGGAGCAGGACTGGGATGGAATAAATGTTGGTACA	1439
Db	77731	GGGCTGTATGTCACCTGTGGAGGAGCAGGACTGGGATGGAATAAATGTTGGTACA	77790
QY	1440	GGAAACGGCAATATACCATCACCAGTGGCCCTGGAAGAGCCTGGTCGAC	1489
Db	77791	GGAAACGGCAATATACCATCACCAGTGGCCCTGGAAGAGCCTGGTCGAC	77840
RESULT 6			
ECEHECCTP			
LOCUS		2407 bp	DNA
			linear
			BCT 12-DEC-1996

DEFINITION	E.coli 0157:H7 katP gene for EHEC-catalase/peroxidase.				
ACCESSION	X89017				
VERSION	X89017.1 GI:1580760				
KEYWORDS	EHEC-catalase/peroxidase; katP gene; Plasmid.				
SOURCE	Escherichia coli.				
ORGANISM	Escherichia coli.				
REFERENCE	Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Escherichia.				
AUTHORS	1 (bases 1 to 2407)				
TITLE	Direct Submission				
JOURNAL	Submitted (22-JUN-1995) W. Brunder, Universitaet Wuerzburg, Institut fuer Hygiene und Mikrobiologie, Josef-Schneider-Strasse 2, D- 97080 Wuerzburg, FRG				
REFERENCE	2 (bases 1 to 2407)				
AUTHORS	Brunder, W., Schmidt, H. and Karch, H.				
TITLE	KatP, a novel catalase-peroxidase encoded by the large plasmid of enterohaemorrhagic Escherichia coli 0157:H7				
JOURNAL	Microbiology 142 (Pt 11), 3305-3315 (1996)				
MEDLINE	97124214				
PUBMED	8969527				
FEATURES	Location/Qualifiers				
source	1..2407				
	/organism="Escherichia coli"				
	/strain="EDL933"				
	/sub_species="serotype 0157:H7"				
	/db_xref="taxon:562"				
	/cell_line="plasmid p0157"				
	119..122				
RBS	131..2341				
gene	/gene="katP"				
	131..2341				
CDS	/gene="katP"				
	/codon_start=1				
	/transl_table=11				
	/product="EHEC-catalase/peroxidase"				
	/protein_id="CAA61429.1"				
	/db_xref="GI:1580761"				
	/db_xref="SPTREMBL:P77038"				
	/translation="MIKKTLPVLILALSGSPSTAVADKKETQFYYPETLDLTPLRLHSPSNPWGADFDYATRFQOLDMEALKDKIDLLTSDWMPADYGHVGPFFIRMAVHCATYRYDGRGGASGGQOREPLNSWPDNVNLDKARLLWPVKKYKSSISWGDLMVUTGNVLESWGFKTLGFPAGREDDESOLVYWDGPNKADNRNKGKLPKLAATQMGIIYVNEGPGGKPLDASAKDIREAFSRMAMDEETVALIAGHTFGKHAASPEKCIGAPGAPVEQGLGKWKNCGTGKTYITSGLEGAWSTSPQFTMQYLNKLYKYEWEHLKSPAGAYQWPKAAITVOADHOPSVLHMFMTDIALKYDPPYKKTITREFLNDKAFEQAFARAFWEKLTDRMGPAARYLGNEVPAESFTWQDPLPADYTMIDGDIKSLKEQVMDLGIPEASELKTAWASASTFRVTDYRGNGNGARIRLQPEINWEVPEKIKKYIASITSLQREFNKKQSGKVSILADLIVLSGNAIEDAARKAGVELEIPEPTPGTNDASQEQTDVASFSVLEPTADGFRNYYSKRSKSHISPVESLIDKASQDLQVPTPTALIGGLRVMDINTNNSLGVFTDTPGVLDNKKFFVNLDMSTRWKSKADKEDTYNGFDRKTGALKWKASSVDLIFSNPELRAVAEYVSDARNKTIHDFVKSWNKVMNSDRFDLNNK"				
sig_peptide	131..199				
	/gene="katP"				
BASE COUNT	691 a	480 c	607 g	629 t	
ORIGIN					
Query Match	72.6%; Score 1081.4; DB 1; Length 2407;				
Best Local Similarity	99.9%; Pred. No. 1e-311;				
Matches 1082; Conservative	0; Mismatches	1; Indels	0; Gaps	0;	
QY	407 GTTGGCGCTTGAACCCCTTAGTATTTTGTCTGTAGTATCTATCCCAGCAATAGGTAT	466			
Db	1 GTTGGCGCTTGAACCCCTTAGTATTTTGTCTGTAGTATCTATCCCAGCAATAGGTAT	60			
QY	467 ATCTGTTCATCAATAAAGTTGACTTTTGTATACAAACATGCGAAATTCCTTTAATCCGG	526			
Db	61 ATCTGTTCATCAATAAAGTTGACTTTTGTATACAAACATGCGAAATTCCTTTAATCCGG	120			
QY	527 AGCTATTCGTATGATAAAAAAACTCTTCTGTTCTGATTTCTTCTGGCGCTATCGGGAG	586			
Db	121 AGCTATTCGTATGATAAAAAAACTCTTCTGTTCTGATTTCTTCTGGCGCTATCGGGAG	180			

QY	587	CTTTTCTACCGCTAGCCGCTGATATAAAGAGACTCAAAATTTCTACTATACCAAAAC	646
Db	181	CTTTTCTACCGCTAGCCGCTGATATAAAGAGACTCAAAATTTCTACTATACCAAAAC	240
QY	647	ACTGGATTAACTCTCTGAGATTACACAGCCCTGRATCAAAATCCCTGGGGGCTGATTT	706
Db	241	ACTGGATTAACTCTCTGAGATTACACAGCCCTGRATCAAAATCCCTGGGGGCTGATTT	300
QY	707	TGATTATGCCACCAAGATTTCACACAGCTGGATATGGAGGCTCTCAAAAAGATATCAAGA	766
Db	301	TGATTATGCCACCAAGATTTCACACAGCTGGATATGGAGGCTCTCAAAAAGATATCAAGA	360
QY	767	TTTCTGACAACTTCCACGAGATTGGTGCCCTCGGATATATGGTCAATATGGTCTTCTT	826
Db	361	TTTCTGACAACTTCCACGAGATTGGTGCCCTCGGATATATGGTCAATATGGTCTTCTT	420
QY	827	TATTCTGATGCTTGGCAGCGTCCGGAACATACAGACATATGATGGCCGGAGGCGC	886
Db	421	TATTCTGATGCTTGGCAGCGTCCGGAACATACAGACATATGATGGCCGGAGGCGC	480
QY	887	CAGTGGTGGTCAGCAACGTTTGAACCGCTTGAACAGCTGGCCGATATAGCTTAACTGGA	946
Db	481	CAGTGGTGGTCAGCAACGTTTGAACCGCTTGAACAGCTGGCCGATATAGCTTAACTGGA	540
QY	947	TAAAGCCGCTGATTTGCTGGCCAGTCAAGAAAAATACGGCTCCAGTATTTCTTGGGG	1006
Db	541	TAAAGCCGCTGATTTGCTGGCCAGTCAAGAAAAATACGGCTCCAGTATTTCTTGGGG	600
QY	1007	AGACCTGATGCTCTGACTGATGATGTCCTGATGATGATGATGATGATGATGATGATG	1066
Db	601	AGACCTGATGCTCTGACTGATGATGTCCTGATGATGATGATGATGATGATGATGATG	560
QY	1067	ATTTCCTGGCGGAGAGAGATGACTGGGAGTGGACCTGGTATATCTGGGGGCTGACAA	1126
Db	661	ATTTCCTGGCGGAGAGAGATGACTGGGAGTGGACCTGGTATATCTGGGGGCTGACAA	720
QY	1127	CAAGCCCTTTCAGATACCGGATATAACCGGAACTTCAGAAACCTTTCGCCGCCAC	1186
Db	721	CAAGCCCTTTCAGATACCGGATATAACCGGAACTTCAGAAACCTTTCGCCGCCAC	780
QY	1187	GCAGATGGGACTTATTTATGCTCAATCTGTAAGCCCGCTGGAAACACAGACTCTCTGC	1246
Db	781	GCAGATGGGACTTATTTATGCTCAATCTGTAAGCCCGCTGGAAACACAGACTCTCTGC	840
QY	1247	TTCCGCGAAGATATCAGGAAGCTTTTACATATGATGGCCATGGATGATGAGAGACTGT	1306
Db	841	TTCCGCGAAGATATCAGGAAGCTTTTACATATGATGGCCATGGATGATGAGAGACTGT	900
QY	1307	GGCCTGATCGCGGAGGCGATACATTTGTAAGCACATGTCAGCGCTCTCCTCAAAA	1366
Db	901	GGCCTGATCGCGGAGGCGATACATTTGTAAGCACATGTCAGCGCTCTCCTCAAAA	960
QY	1367	ATGATTGGCGCAGGCGCTGATGGTACCTGTGGAGGAGCAGGACTGGGATGGA AAA	1426
Db	961	ATGATTGGCGCAGGCGCTGATGGTACCTGTGGAGGAGCAGGACTGGGATGGA AAA	1020
QY	1427	TAAATGGTACAGGAACGGCAATATACCATCAGCTGGCTGGAGGAGCCTGGTC	1486
Db	1021	TAAATGGTACAGGAACGGCAATATACCATCAGCTGGCTGGAGGAGCCTGGTC	1080
QY	1487	GAC 1489	
Db	1081	GAC 1083	

RESULT 7
AE013690/c

LOCUS AE013690 12268 bp DNA linear BCT 26-JUL-2002
DEFINITION Yersinia pestis KIM section 90 of 415 of the complete genome.
ACCESSION AE013690 AE009952
VERSION AE013690.1 GI:21957603
KEYWORDS Yersinia pestis KIM.
SOURCE

ORGANISM

Yersinia pestis KIM
Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
Yersinia.

REFERENCE
AUTHORS

1 (bases 1 to 12268)
Deng, W., Burland, V., Plunkett, G. III, Boutin, A., Mayhew, G. F.,
Liss, P., Perna, N. T., Rose, D. J., Mau, B., Zhou, S., Schwartz, D. C.,
Fetherston, J. D., Lindler, L. E., Brubaker, R. R., Plana, G. V.,
Stratley, S. C., McDonough, K. A., Nilles, M. L., Matson, J. S.,
Blattner, F. R. and Perry, R. D.
Genome Sequence of Yersinia pestis KIM
J. Bacteriol. 184 (16), 4601-4611 (2002)

TITLE
JOURNAL

Genome Sequence of Yersinia pestis KIM
J. Bacteriol. 184 (16), 4601-4611 (2002)

PUBMED
REFERENCE
AUTHORS

2 (bases 1 to 12268)
Deng, W., Burland, V., Plunkett, G. III, Boutin, A., Mayhew, G. F.,
Liss, P., Perna, N. T., Rose, D. J., Mau, B., Zhou, S., Schwartz, D. C.,
Fetherston, J. D., Lindler, L. E., Brubaker, R. R., Plana, G. V.,
Stratley, S. C., McDonough, K. A., Nilles, M. L., Matson, J. S.,
Blattner, F. R. and Perry, R. D.
Direct Submission
Submitted (21-FEB-2002) Genetics, University of Wisconsin, 445
Henry Mall, Madison, WI 53706, USA

TITLE
JOURNAL

Genetics, University of Wisconsin, 445
Henry Mall, Madison, WI 53706, USA

FEATURES
source

1..12268
Location/Qualifiers
/organism="Yersinia pestis KIM"
/strain="KIM"
/db_xref="taxon:187410"
complement(110..2365)
/gene="katG"
/note="y0870"
complement(110..2365)
/gene="katG"
/function="enzyme: detoxification"
/note="Residues 37 to 748 of 751 are 55.68 pct identical
to residues 26 to 725 of 726 from E. coli K12: B3942;
residues 15 to 751 of 751 are 100.00 pct identical to
residues 1 to 737 of 737 from GenPept :
>gI|AAD37313.1|AF135170_4 (AF135170) catalase/peroxidase
[Yersinia pestis]"
/codon_start=1
/transl_table=11
/product="catalase; hydroperoxidase HPI (I)"
/protein_id="AA084455.1"
/db_xref="GI:21957604"

/translation="MLKNSCLVKGDDLYMLKLPVLITLAIHVHTPTAWAARPKTID
SFYLPKSLDLSPLRHNTESNPKYQDFNAQOFTLDEAVKDKDKVLTVTSOOWWPA
DYNGYPPFFIRMAWHGAGTIRYDGRGAGDQOQFEPPLNSWPDNADKARLLWPI
KKYKAKISWGLDMLVGNVALESNGFKTLGAFAGREDDQSDLYWAGNKLSDNR
DKNGKLPPLAATOMGLIYVNPESGKPPDPAAKKIDIRAFARMANDEETVALLTG
GHTGKAHGAASPEKCLGAAPGAGLQSGQGLGWANKCGNGDKITITSGLEGATTTDP
THFTMQYLSNLYKHEWVLTSPAGAWQPKNAANVYPDATDPTKPHLMFTDIAL
KVDPYKKITTRFLENPEEFKMAFARAFWKLTTHRMGMPAARYLGDEVPKFTFWQDPL
PAANYKIDSADISLKDILKTLGSDTKLITAWASATFRGTDPRGGDNGARILRLA
PKQWPVNDPAELHSLVAALMEVQNNFNKDRSDGKVSLSDLTVLIGNAAIEDAAKKA
GYSISIPPTGRTDASOEDTVDSFVLEPTADGFRNYDAKNTLSPIASLIDRANK
LELYPENTVLIIGLRVLDVNSGSKAGVLTNPQUNNFFVNLMDTKWTKSPKA
EGYFDGDRKTKLTKWTASSVDLVFGSNPELRAVAEVIYASDDAKEFEVHDTTKWEKV
MNLDRFDIKNN"
2811..3800
/gene="y0871"
2811..3800
/gene="y0871"
/function="putative transport"
/note="sugar (D-ribose?) transport; residues 6 to 329 of
329 are 100.00 pct identical to residues 1 to 324 of 324
from GenPept : >gI|AAD37312.1|AF135170_3 (AF135170)
D-ribose binding protein [Yersinia pestis]"
/codon_start=1
/transl_table=11
/product="putative
ABC transporter"
/protein_id="AA084456.1"
/db_xref="GI:21957605"
/translation="MNINLMVNIYSMRKIMKNKNTLLKSLCTAALISMAGVAGAAASNGL

gene

2811..3800
/gene="y0871"
2811..3800
/gene="y0871"

CDS

2811..3800
/gene="y0871"
/function="putative transport"
/note="sugar (D-ribose?) transport; residues 6 to 329 of
329 are 100.00 pct identical to residues 1 to 324 of 324
from GenPept : >gI|AAD37312.1|AF135170_3 (AF135170)
D-ribose binding protein [Yersinia pestis]"
/codon_start=1
/transl_table=11
/product="putative
ABC transporter"
/protein_id="AA084456.1"
/db_xref="GI:21957605"
/translation="MNINLMVNIYSMRKIMKNKNTLLKSLCTAALISMAGVAGAAASNGL

QY 729 CAGCTGGATAGGAGCTCTGAAAAAGATATCAAGATTGTGCTGACAACTTCCCAAGAT 788
 Db 2131 ACATGGATCTGAAGCAGTAAGAAAGATATAAAACAGTTCTTACCACATCACAGAC 2072
 QY 789 TGGTCCCTCGGATATGTTCAATATGTTCTTCTTTATTCGTATGCTTGGCAGCGT 848
 Db 2071 TGGTGGCTCGTATATGTTCAATATGTTCTTCTTTATTCGTATGCTTGGCAGCGT 2012
 QY 849 GCCGGAACATACAGACATATGATGCGCGGGAGGCCAGTGGTGTGACAACTGTTT 908
 Db 2011 CGCGGAACCTACCATATATGATGCGCGGTGGTGGTCCGATGCGCGGACAAAGATT 1952
 QY 909 GAACCGCTGAACAGCTGCGCGGATAACGTTAATCTGATAAAGCCCGCTGATTCGCTGG 968
 Db 1951 GAGCCACTCATAGCTGCGCAGATAACGCCAACCTTGTAAAGCGCGTCTGCTGG 1892
 QY 969 CGAGTCAAGAAAAATACGGCTCCAGTATTTCTTGGGGAGACCTGATGCTTGGACTGGT 1028
 Db 1891 CCTATAAAAAAATATGCGGCTTAAATATCTGCGGCGACCTGATGCTTACAGGC 1832
 QY 1029 AATGTTGCCCTGAATCCATCGATTTAAACGCTGGGATTTGCTGCGGAAGAGAGAT 1088
 Db 1831 AATGTCGCGCTGAATCTATGCGGTTTAAACGCTGGGTTGCGAGGCGCTGAGGAT 1772
 QY 1089 GACTGGAGTCGGACCTCGTATCTACTGCGGCGCTGACAAACGCTCTTGCAGATAACCG 1148
 Db 1771 GATTGCAATCGATCTGCTCTACTTGGGGCGCAGCAACAGATGCTGCTGATAACAG 1712
 QY 1149 GATAAAACGGGAACCTTCAGAAACCTCTTCCGCGCAGCAGATGGGATTTATTTATGTC 1208
 Db 1711 GATNAATATGCAAGCTACCAACCGCTGCGGCGCACAAATGGGGTTGATTTAGCTA 1652
 QY 1209 AATCTGGAAGCGCCGCTGAAACACAGATCTCTGCTTCCGCGAAGATATCAGGAA 1268
 Db 1651 AACCTGGAAGCAAAATGTTAAACACAGCCGCTGCTGCGCAAAAGATATTCGTAA 1592
 QY 1269 GCTTTTACGATGATGCGATGATGATGAGGAGCTGTCGCTGATGCGGAGGAGCAT 1328
 Db 1591 GCGTTTGCAGATGCAATGAATGATGAGGAGACCGCTGGCATTTGCGGGAGGCGAT 1532
 QY 1329 ACATTTGTTAAAGCAGATGTCAGCGCTCTCTGAAATATGTTGCGCAGGCGCTGAT 1388
 Db 1531 ACCTTGGTTAAAGCGATGCGCGCAGCGCTCTCTGAAATATGCTTGGTGGCAGCAGCG 1472
 QY 1389 GTTGCACCTGTGGAGGAGCAGGAGCTGCGGATGGAATAAATATGTTGACAGGAACGCG 1448
 Db 1471 GAAGCGGCTAGAACAAACAGGTTTAGATGGCGGAATAAATGTTGCTTACGTAATGTT 1412
 QY 1449 AATATACCATCAGCAGTGGCTGGAAGGAGCCTGCTGAC 1489
 Db 1411 AAGGACAGGATAACCAAGTGGTTGGAAGGCGCATGGCAAC 1371

RESULT 8
 AF135170
 LOCUS
 DEFINITION
 Versinia pestis ribose ABC transporter (rbsA) gene, partial cds;
 D-ribose binding protein (rbsB), catalase/peroxidase (katY),
 cytochrome c (cybC), cytochrome b (cybB), dimethyl sulfoxide
 reductase subunit C (dmsC), dimethyl sulfoxide reductase subunit B
 (dmsB), dimethyl sulfoxide reductase subunit A (dmsA), and
 L-ribulose-phosphate 4-epimerase (arad) genes, complete cds; and
 unknown genes.
 AF135170
 AF135170.1 GI:5002118
 ORGANISM
 Versinia pestis.
 Versinia pestis
 Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 Yersinia.
 1 (bases 1 to 13632)
 Garcia,E., Nedialkov,Y.A., Elliott,J., Motin,V.L. and Brubaker,R.R.
 Molecular characterization of KatY (antigen 5), a thermoregulated

chromosomally encoded catalase-peroxidase of Yersinia pestis
 J. Bacteriol. 181 (10), 3114-3122 (1999)
 99255535
 10322012
 2 (bases 1 to 13632)
 Garcia,E., Nedialkov,Y.A., Elliott,J., Motin,V.L. and Brubaker,R.R.
 Direct Submission
 Submitted (16-MAR-1999) BBRP, Lawrence Livermore National Lab, 7000
 East Ave., Livermore, CA 94550, USA
 Location/Qualifiers
 1..13632
 /organism="Yersinia pestis"
 /strain="KIM"
 /db_xref="taxon:632"
 complement(<1..1034)
 /gene="rbsA"
 complement(<1..1034)
 /gene="rbsA"
 /note="ATP-binding protein"
 /codon_start=1
 /transl_table=11
 /product="ribose ABC transporter"
 /protein_id="AAD37310.1"
 /db_xref="GI:5002119"
 /translation="MNRSPDIVRAEDISMREPGTALDLSVNYTVYRGKVNVIIGENG
 AKSTIMKILAGVQOQTSGQIYLNGKYSIANTREAAALGIGMVOELNLSENINVAE
 NIFLREIOHGLKPIQAQERVTQMLVRLDQVISPQEKVSNLKYGQQQLIEAKAL
 ASQADILIDELTSALETFDVLFRILTROGVSIVYISHRLLEMAIGDSITIL
 RQGRQAFAEKVIDIPWITREMLGSDPVSFLRPGRTFGPMEVEDVTLISGNT
 VNVHSLVHRAGEIVIGIYLMGAGRTLFECILGTLQNYLTGLTVRLNGTVMNAKTSAE
 RIRFGMSLVP"
 complement(1034..1651)
 /note="ORF2; hypothetical protein"
 /codon_start=1
 /evidence=not_experimental
 /transl_table=11
 /product="unknown"
 /protein_id="AAD37311.1"
 /db_xref="GI:5002120"
 /translation="MRFRAISKAGALLMISAVLVACTVVDLDENKPILPVDPNAVVS
 DYNQTPKVAATIAKWNPFANSAISQWQVQKQSQPAVKNNQSRFRTGKVA
 VETEGELIRDAIDGEOVLQIGPIVGNALIRDASTFIRDFDQNOVQYQALAKALS
 LRQEVAKPDAISWIGQVEVLAATLTPTLSLNGVPLSLNKESH"
 complement(1744..2718)
 /gene="rbsB"
 complement(1744..2718)
 /gene="rbsB"
 /codon_start=1
 /transl_table=11
 /product="D-ribose binding protein"
 /protein_id="AAD37312.1"
 /db_xref="GI:5002121"
 /translation="MVNIYSMRKIMKNILLKSLTAALISMAGVAGASNLIAIIT
 PSHDNPFKAAEAGAKAKATGELTYTLVAHSHDDVNKQNLQIETAIARAKAILIDNA
 GSDATIGPLKKAAGAPAFIDIREINETGIAVSQIVSNYQAGQIGAEKFTYLMGGK
 PGTIELGREDTNHVRQSGHDVIDEHSDEKMKVAQQTANNSQTFEAFNMESIIQAN
 PDITGVLSGNDTNALGAEALKAAGNDVVVVGFDGSDYVRDSIINNGNIKATVLQPG
 WAAQAMVAVQADQLYTKTGLEEKQMLDCVLIDENNAKLNLFALDK"
 3221..5434
 /gene="katY"
 3221..5434
 /gene="katY"
 /note="antigen 5; katP homolog"
 /codon_start=1
 /transl_table=11
 /product="catalase/peroxidase"
 /protein_id="AAD37313.1"
 /db_xref="GI:5002122"
 /translation="MLKKILPVLITLAIIVHTPTAWAAEAPKTDSTFLPKSLDLSPLR
 LHNIESNPYKDNVYACQFTLDLEAVKKDKITVLTSDQWPDADYGNYPFFIRMAW
 HGATYRIYRGAGGAGCGQORFELNSPDNANLDKARLLWFIKKYGNKISWGLIM
 VLTGNVALESRGFKTLGFAAGREDWQSDLVYTWGAGNKMLSDNRDKNGLKPLPLAATQ
 MGLIYVNEPGNPKDPVAAAKDIREAFARMANDEETVALIAGGHTFGKAHGAASPE

```

KCLGAPGAGLEQOGLGWANKCGSGNGKDTITSGLEGAWTTDPTHTMQYLSNLYKH
EWLYTSFAGAWMPKNAANVVPDADPTKEHPLKMFETDIALKVDPEYKKTITREFL
ENPEFKAAPKAPKALTRHDGPAARYLGDVEPKETFTWQDPLPAANYKMTSDAIS
EKDKLITKLTGSDTLIKITAWASASTFRGTDFRGDNGARIRLAPKQWMPVPAELH
SVLAALMEYQNNFNDRSGKVSLSDLIVLGGNAIEOAAKAGYSISIPPTPGRTD
ASQETDVSSEAVLEPTADGFRNYDYAKRNTLSPASLIDRANKLELVPEWTLIGG
LRVLDSNGSGKAGVLTWTPGGLNNFVLLDMSTKTKSPKAGYFDGYDRKTKGL
KWTASVDLVFNPFLRAVAEVIYASDDAKEAFVHDTFKVWEKVMNLDKRFKNN"
5511..5894
/gene="cybc"
CDS
5511..5894
/gene="cybc"
/gene="cybc"
/note="similar to Escherichia coli cytochrome c-562"
/codon_start=1
/transl_table=11
/product="cytochrome c"
/protein_id="AAD37314.1"
/db_xref="GI:5002123"
/translation="MRKPIIAGVPSLLITSTFAASLQDDNMILIANIGIVSSSTDPT
KYTISLEKRRNALDAQAIIPKLEGAEDSPEDIRHGFDFLLIEQIDKTKQWAE
GNIQEVKSVGVINIRNTYHSRY"
6042..6572
/gene="cybb"
CDS
6042..6572
/gene="cybb"
/note="similar to Escherichia coli cytochrome b-561"
/codon_start=1
/transl_table=11
/product="cytochrome b"
/protein_id="AAD37315.1"
/db_xref="GI:5002124"
/translation="MHKYSCLQIGHWLVLLVIANSSIELRGAFAPRSYOPWKKMI
HSCGIALVLMFTLILQRYPTPIVPPKSPMIVGLAHVGHWYILFLIALPIGI
AIIYCRGSGIAFGLIPLHAEQAFDLADTLKAYHLLANMSYFVIGLHAAALLHY
VLDNTLLRMPEKRG"
complement(6829..7539)
/note="ORF7: hypothetical protein"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="unknown"
/db_xref="GI:5002125"
/translation="WIKDVTYQDIVKYDYQDSYQYKIGITGRVLGALFYCEPDSA
ECRDIVQURNTWAEWPGYGADELMPATILAAQSOYOSIAHQSAAPLAETLE
EAWRLFIGPALPAPPQSVYLDKMDVLFSDSTLKRNMQORHVEVTLKQOEPEDH
FGLLMMMAWLAEPQFVDPLVLLADHELLPWSYRYLALLQSDAGHPFYQGLAQLTIL
AHWQHELOVTPAGVELYR"
complement(7685..8545)
/gene="dmsC"
CDS
complement(7685..8545)
/note="dmsC"
/note="similar to Escherichia coli DMSO reductase"
/codon_start=1
/transl_table=11
/product="dimethyl sulfoxide reductase subunit C"
/db_xref="GI:5002126"
/translation="MGSGHWEPLMVFVLGCVVGGFVLGLALIFGLSGRQORV
HRSFVLYWLMALFASLTHLGSMPAFNSLIRGESALSNEIAGSFFAVAGFYW
LVAIDKPAAGKILWVYVGLVYFYVANCYVSINTVPTWDSIYPLGVLITLI
GGPLGILLQLADVNGRAMQLPMISVLIIISVASVIMQASLPLIYSYQOASAL
IPDYGLSMVWRLVILVGLGCWICPLIRGRRPMTLVWVLGTLVLAGELIGRGVYGL
HMTVGMVAGG"
complement(8547..9164)
/gene="dmsB"
CDS
complement(8547..9164)
/note="dmsB"
/note="similar to Escherichia coli DMSO reductase"
/codon_start=1
/transl_table=11
/product="dimethyl sulfoxide reductase subunit B"
/protein_id="AAD37318.1"

```

```

/db_xref="GI:5002127"
/translation="MTQYGFYIDSRCTGCKTCELACKDFKLNLPDVSFRRIYAYG
GWQDNGTWHQNTFYAYTYSIACNHCSDPACTKVPSPGAMHNRDDGQFVVVNDICIGC
RYCHNACYPAPDAEKGHMTKDCQCYERVAVGKPKICVDSCPIRALDMPIDELRE
KYGDIAETAPLPAAHFTLPIVILKNANSPRPVDTTGHLANPKEV"
complement(9176..11647)
/gene="dmsA"
CDS
complement(9176..11647)
/note="dmsA"
/note="similar to Escherichia coli DMSO reductase"
/codon_start=1
/transl_table=11
/product="dimethyl sulfoxide reductase subunit A"
/protein_id="AAD37319.1"
/db_xref="GI:5002128"
/translation="MOLSNLKMKNESRLTTGVTRRKLVTITLVGGLAMATGAFSL
PFSRTARVQSALNPTSANDKVINSGSRCPPLRMHVADGEIKYVEDNTGDD
DREGLHOVRACLGRSMRRRYNPDRLKYPMKRIGARGEGKFRKISWEAFETAGSM
QNIIRKEYGNEAIYLNVTGTTIGTTRSPWPGATLLARLMCCGGLNYGDIYSTAOI
AAGLNYTGGWAGNDSPIENSLVLFNPNPAETRMSSGGVYLYLEQAREKSNAM
IVIDPRYTDAAGREDEWIPRPGTDAALASALAYVMTENLVDPFLLQYCVGVDEK
TLPEGAPENSHYKAYILGOCADTKATPOWAAITGIPAAKITTLGREGISGVKPAYIS
QCGPQRHSGENTSRATJAMLAILTCNVGINGSNGAREGSYSLPVRMPTPENPVKT
SISMFELTDAIERGPEMTATRGVQKDLDPYRIFWYASNCLYNQHSLENRTHDI
LQDDKCEMIVVDNMTSSAETADILIDPCTASEMDFCLDASSGNMAYVIFADQVI
KPRFECKNIYDITSLARHMGVEKFTGRTQEBWHRHLYQOQSAIPELPSFEDFPA
PLPVYAAGFESHVDPLIEQYPLQLTGFHYKARTSTYGNVDVKASCLQEMINPMDA
ERRGISNGDMVIRFNGRGVQVONAKVTPRMGVVALGEGAWYSPDAKRIDRAGSINV
LTQRPSPLAGKNPSHTNLVQVAKA"
complement(12089..12859)
/gene="arad"
CDS
complement(12089..12859)
/note="arad"
/note="similar to Escherichia coli ribulose epimerase"
/codon_start=1

```

```

Query Match      36.0%; Score 536.6; DB 1; Length 13632;
Best Local Similarity 69.7%; Pred. No. 1.5e-148;
Matches 740; Conservative 0; Mismatches 319; Indels 2; Gaps 1;

QY 429 TATTTTTCGTCTAGTATCTATCCAGCAAGTAGGTATATCTCTGTCATCAATAAGTT 488
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 3115 TATCAATAATAGATAGATCTATCTATTAATAGGACTACGAATTTGTTTCCATTT 3174

QY 489 GACTTTTGTATACAAATCGAATTTCCCTTAATCCGAGACTATTCGTATGATAAAAAA 548
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 3175 AATCATGTTAAAAATAGTTGCTTTTACTTAAAGGGGACTTAT--ATATGTTAAAAAA 3232

QY 549 ACTCTTCTCTGTTCTGATTTCTTGGCGGTATCGGGGAGCTTTTCTACCGCTGTAGCCGT 608
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 3233 ATCTTACCCTACTAATACTCTCGCCATTGTACATAATACACCTACGCTTGGCGAGCA 3292

QY 609 GATAAAAAGAGACTCAAAATTTCTACTATCCAGAACACTGGATTAATCTCTCTGAGA 668
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 3293 GAAGACCCAGACTGATAGTTTACTTACCCAAAAGAGCTGATCTCTCCACTCCGT 3352

QY 669 TTACACAGCCCTGAATCAATCCCTGGGGGCTGATTTTGAATATGCCACAGATTCAA 728
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 3353 TTACATAATAGAGTCCATCTCTACGGAAGAGGATTTTANTATGCTCAACAGTTTAA 3412

QY 729 CAGCTGGATATGAGGCTCTGAAAAAGATATCAAGATTTGCTGACAACTTCCCAAGAT 788
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 3413 ACACCTGGATCTTGAAGCAGTAAAGAAAGATATAAAACAGTCTTACCACATCACAGAC 3472

QY 789 TGGTCCCTCGCGGATATGTCATATGCTCTCTTCTTTTATTCGTATGCTTGGCAGGT 848
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 3473 TGGTCCCTCGCTGATGATGATGTTATGTTCCATCTTTTATTGCTGATGCTGGCATGT 3532

QY 849 GCGGGAACATACAGGACATATGATGCCGGGAGCGCCAGTGGTGGTGGTGGGCGCAAGATTT 908
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 3533 GCGGGAACCTACCGCATATATGATGCCCTGGTGGTGGTGGGCGGCGCAAGATTT 3592

```

QY	909	GAACCGCTGAACAGCTGCGCGGATACGTTAACTTGGATGAAGACCGCTGATGCTGTGG	968
Db	3593	GAAGCACTCATAGCTGCCAGATACCGCAACTGTATTAACCGCTGCGTTCGTGG	3652
QY	969	CCAGTCAAGAAAAAATACGCTCCAGTATTTCTGGGAGACCTGATGTCCTGACTGT	1028
Db	3653	CCTATAAAAAAATATGCGCTTAAATATCGTGGGCGCACCTGATGTCCTTACAGGC	3712
QY	1029	AATGTTGCCCTGAATCATCGTGGATTTAAACGCTGGATTTCTGCGGAGAGAGAT	1088
Db	3713	AATGCGCGCTCGAATCATGCGGTTTAAACGCTGGGTTTTCAGAGGCGCGTGAGGAT	3772
QY	1089	GACTGGGAGTCCGACCTGCTATCTGGGGCCTGACAAACAGCCTCTTCAGATACACCG	1148
Db	3773	GATGGCAATCCGATCTGCTACTGCGGCGAGGCAACAAGATGCTGTCTGATACAGG	3832
QY	1149	GATAAACCGGAACTTCAGAACCTTTCGCCGCCACGAGATGGACATTTATGTC	1208
Db	3833	GATAAAATGGCAAGCTACCCAAACCGCTGCGGCGACACAAATGGGGTTGATTTACGTA	3892
QY	1209	RATCTGTAAGCCCGCGTGGAAACACAGATCCTCTGCTCCGCGGAAGATATCAGGGA	1268
Db	3893	RACCTGTAAGCACCAATATGGTAACACAGACCCGCTTCTGCGGCAAAAGATATTCGTGA	3952
QY	1269	GCTTTTTCAGCTATGCGCATGATGATGAGAGACTGTGGCCCTGATCGCGGAGGCGAT	1328
Db	3953	CGGTTTTCAGGATGGCAATGAATGATGAGGAGACCGCTGGCATTGATTGGGAGGCGAT	4012
QY	1329	ACATTGTTAAGACATGTTGACGCTCTCCTGAAATGATTTGCGGAGGCGCTGAT	1388
Db	4013	ACCTTTCGTTAAGCCATGGCGGACGCTCTCCTGAAATGATTTGCGGAGGCGCTGAT	4072
QY	1389	GCTGACCTGTGGAGGAGCAGGACTGGGATGGGAAATAATGTTGTCAGCAACAGGC	1448
Db	4073	GAAGCGGGCTAGAACACAGGTTTAGGATGGCGAATNAATGTTGTCAGGTAATGTT	4132
QY	1449	AAATATACATCACCAGTGGCTTGAAGAGCCTGGTGCAC	1489
Db	4133	AAGGACACGATAACAGTGGGTTGGAAGGGCATGGCAAC	4173
RESULT	9		
LOCUS	AJ414156	220050 bp	DNA linear BCT 06-JUN-2002
DEFINITION	Yersinia pestis strain Co92 complete genome; segment 16/20.		
ACCESSION	AJ414156	AL590842	
VERSION	AJ414156.1	GI:15981150	
KEYWORDS	Yersinia pestis.		
SOURCE	Yersinia pestis.		
ORGANISM	Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Yersinia.		
REFERENCE	1 (bases 1 to 220050)		
AUTHORS	Parkhill, J., Wren, B.W., Thomson, N.R., Titball, R.W., Holden, M.T.G., Prentice, M.B., Sebahia, M., James, K.D., Churcher, C., Mungall, K.L., Baker, S., Basham, D., Bentley, S.D., Brooks, K., Cerdeno-Tarraga, A.M., Chillingworth, T., Cronin, A., Davies, R.M., Davis, P., Dougan, G., Feitwell, T., Hamlin, N., Holroyd, S., Jagels, K., Leather, S., Kariyeh, A.V., Moule, S., Oyston, P.C.F., Quail, M., Rutherford, K., Simmonds, M., Skelton, J., Stevens, K., Whitehead, S. and Barrell, B.G.		
TITLE	Genome sequence of Yersinia pestis, the causative agent of plague		
MEDLINE	Nature 413 (6855), 523-527 (2001)		
PUBLISHED	21470413		
REFERENCE	1586360		
AUTHORS	2 (bases 1 to 220050)		
TITLE	Direct Submission		
JOURNAL	Submitted (04-OCT-2001) Submitted on behalf of the Yersinia sequencing team, Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA E-mail: parkhill@sanger.ac.uk		
COMMENT	Notes: Details of Y. pestis sequencing at the Sanger Centre are available on the World Wide Web.		

FEATURES	(URL, http://www.sanger.ac.uk/Projects/Y_pestis/).
source	Location/Qualifiers
	1. 220050
	/organism="Yersinia pestis"
	/strain="CO92"
	/db_xref="taxon:632"
gene	/note="biovar: Orientalis"
	complement(191..715)
	/gene="YPO3215"
	/note="arol"
CDS	complement(191..715)
	/gene="YPO3215"
	/EC_number="2.7.1.71"
	/note="Similar to Escherichia coli shikimate kinase II AROL SW:AROL_ECOLI (P08329) (174 aa) fasta scores: E(): 1.5e-31, 55.4% id in 168 aa, and to Erwinia chrysanthemi shikimate kinase Arol or AROM SW:AROL_ERWCH (P10880) (173 aa) fasta scores: E(): 0, 63.7% id in 168 aa"
	/codon_start=1
	/transl_table=11
	/product="shikimate kinase II"
	/protein_id="CAC92450.1"
	/db_xref="GI:15981151"
	/db_xref="SPTREMBL:O8ZC15"
	/translation="MTQTIFMVVGARGAGKTTIGKALQAOLGVRFRVDTDLFMQQTSMQT VAEVYSGWDGFRLRSMALQAVTAPKTVATGGGAVLSSENRAFMEDHGRVLYLRA SAAVLAKRLAEDPEEAQRPSLTGKRFIVEIILDVLASREALIQDVAAHVLDGTQTPSLV VEQIQLMLTGMVVK"
misc_feature	complement(209..709)
	/gene="YPO3215"
	/note="Pfam match to entry PF01202 SKI, Shikimate kinase, score 245.00, E-value 1.1e-69"
misc_feature	complement(470..544)
	/gene="YPO3215"
	/note="PS01128 Shikimate kinase signature."
misc_feature	complement(668..691)
	/gene="YPO3215"
	/note="PS00017 ATP/GTP-binding site motif A (p-loop)."
gene	complement(962..1519)
	/gene="YPO3216"
CDS	complement(962..1519)
	/gene="YPO3216"
	/note="Similar to Rhizobium meliloti phospholipid N-methyltransferase PmtA TR:AAG10237 (EMBL:AF201699) (200 aa) fasta scores: E(): 1.1e-07, 25.9% id in 185 aa, and to Bradyrhizobium japonicum PmtA protein TR:O9LCT2 (EMBL:Y09633) (199 aa) fasta scores: E(): 0.0057, 24.2% id in 178 aa, and to Klebsiella pneumoniae hypothetical 22.7 kDa protein TR:O87755 (EMBL:AJ011907) (196 aa) fasta scores: E(): 0, 52.9% id in 191 aa"
	/codon_start=1
	/transl_table=11
	/product="putative methyltransferase"
	/protein_id="CAC92451.1"
	/db_xref="GI:15981152"
	/db_xref="SPTREMBL:O8ZC14"
	/translation="MLNIVRNFKQSFYLQRFIAPSPRTVSLVSPSPWLCQAMLNQI DWTQNLNIALGADGVLAKRLSHMSANSLSQAYEIOFPHVHALHQLKDSRLQADR SAROLQEDVDFVCCPLLSIPTKISIRILOQTLQRLRANGVLVLFQYSHLSSELSL RYFNWKKIRVVRNFPFALVYICHPN"
gene	1787..2002
	/gene="YPO3217"
CDS	1787..2002
	/gene="YPO3217"
	/note="no significant database hits"
	/codon_start=1
	/transl_table=11
	/product="hypothetical protein"
	/protein_id="CAC92452.1"
	/db_xref="GI:15981153"
	/db_xref="SPTREMBL:O8ZC13"
	/translation="MIGTDVGIIVVAVNVAETNDRGRSISCHFFTIQSSPNAHPLNVD IKNIASHVLISIVACSIKKAAMRLS"

```

gene
complement(2249..2770)
/gene="YPO3218"
CDS
complement(2249..2770)
/gene="YPO3218"
/notes="Similar to Escherichia coli hypothetical protein
y4fe SW:YAFE_ECOLI (P30866) (207 aa) fasta scores: E(): 0,
65.1% id in 169 aa, and to Pseudomonas aeruginosa y4fe or
PA3119 TR:087011 (EMBL:U93274) (187 aa) fasta scores: E():
0, 55.2% id in 172 aa"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="CAC92453.1"
/db_xref="GI:15981154"
/db_xref="SPTREMBL:O82C12"
/translation="MQAAADKKLTNIEIQGLAESLPFADQSFQIVISRYSAHHWHVDV
GKALREVKLVRPGGVIFMDVSPGHVLDIYLVQVEVLROTSHVRNYSFGWLSLF
TEGLVINEVTSRLYLEFSWVARMRTPAHPATAREFQKLASDGVINHAIQADGS
FTSDIMMIIVARD"
complement(3016..3285)
/gene="YPO3219"
CDS
complement(3016..3285)
/gene="YPO3219"
/notes="no significant database matches"
/codon_start=1
/transl_table=11
/product="putative membrane protein"
/protein_id="CAC92454.1"
/db_xref="GI:15981155"
/db_xref="SPTREMBL:O82C11"
/translation="MRGSMVGISEESIRRLNFKNAVGHSAKVAIISAVFIIMLF
YHLEGLVLAGIGIFVSTTWKKKQKYSKNQIQIEEDKV"
complement(3088..3138)
/gene="YPO3219"
/notes="one of 2 probable transmembrane helices predicted
for YPO3219 by TMHMM2.0"
complement(3151..3216)
/gene="YPO3219"
/notes="one of 2 probable transmembrane helices predicted
for YPO3219 by TMHMM2.0"
complement(3468..4226)
/gene="YPO3220"
complement(3468..4226)
/gene="YPO3220"
/notes="Similar to Escherichia coli hypothetical protein
ycgr SW:YCGR_ECOLI (P76010) (244 aa) fasta scores: E(): 0,
40.7% id in 241 aa"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="CAC92455.1"
/db_xref="GI:15981156"
/db_xref="SPTREMBL:O82C10"
/translation="MGPVSETSKQPVKKNKLAICAILRDLOKNDTAVMVTHARGOFI
SRLLDQPTNDFIDFGSVENENVLALGAEOLTIIVPTGAKIEFTCKLKHVELS
LPAFSAIEPQLFYIQRREYFVSIPQWPAFYCSGFPGTQYKYLADISLGMGLY
AMKSEFPQGSGLVADAAVDICGFLFKLDLOFIRALDKQVNNNGEMLTVQRLSFK
FLKLSPIQKGLQRAIFELEKQQTAKARFQEG"
complement(4554..5813)
/gene="YPO3221"
/notes="proA"
complement(4554..5813)
/gene="YPO3221"
/EC_number="1.2.1.41"
/notes="Similar to Serratia marcescens gamma-glutamyl
phosphate reductase ProA SW:PROA_SERMA (P17857) (417 aa)
fasta scores: E(): 0, 85.1% id in 417 aa, and to
Escherichia coli gamma-glutamyl phosphate reductase ProA
SW:PROA_ECOLI (P07004) (417 aa) fasta scores: E(): 0,
72.4% id in 417 aa"
/codon_start=1
/transl_table=11
/product="gamma-glutamyl phosphate reductase"

```

```

/protein_id="CAC92456.1"
/db_xref="GI:15981157"
/db_xref="SPTREMBL:O82C09"
/translation="MNLRLHMGKAQKASWOLAMLSSTAKKNQALAVIANLLESEOTI
LQANERDMAAARESGHSEALLRLLLTPARLAATANDVQVRLNDPVGRIVDLSLLD
SGIKLERRVPVIGVIGVYEARNVWIDVNSCLKTGNVILRGGEETHYNTQATVNV
IQRLQSGGLPAAAVQAIKESPDQLVNEELRLDRYDMLIPRGASLHKLCREOSTIP
VITGGIGVCHTFVDNADFEKALLVIENAKIQRPSACNSLETLLVHQAVAKTEPLLS
ARMHAFGVTLHASPILAMPYADKAVVAEADYDDEWLSLDLVNDIVTIDDAIDH
IREHGTSHSDAILTRLSHAEYFVRAVDSSAVVYNASTRTFDGQFGIGAEVAVSTOK
LHARGPMGLDALTYKWIQVGGDLVRS"
complement(4572..5744)
/gene="YPO3221"
/notes="Pfam match to entry PF00171 aldedh, Aldehyde
dehydrogenase family, score -217.90, E-value 0.00013"
complement(4776..4841)
/gene="YPO3221"
/notes="PS01223 Gamma-glutamyl phosphate reductase
signature"
complement(5823..6926)
/gene="YPO3222"
/notes="proB"
complement(5823..6926)
/gene="YPO3222"

Query Match      36.0%; Score 536.6; DB 1; Length 220050;
Best Local Similarity 69.7%; Pred. No. 3e-148;
Matches 740; Conservative 0; Mismatches 319; Indels 2; Gaps 1;

QY 429 TATTTTGTCTGTAGTACTATCCAGCAATAGGTATATCTCTGTGTCATCAATAAGTT 488
DB 121630 TATCAATAATAGATGAATCTATATTAATAGGACTACGAATGATATTTTCCATTT 121689

QY 489 GACTTTGTATACAAATCGGAATTTCCCTTAATCCGGAGCTATTCGTATGATAAAAAA 548
DB 121690 AATCATGTTAAAAAATAGTTCCTTTTAGTTAAGGGGACTTAT--ATATGTTAAAAA 121747

QY 549 ACTCTTCTCTGTCTGTATCTCTCTGGCGCTATCGGGAGCTTTTCTACCGCTGTAGCGCT 608
DB 121748 ATCTTACCGCTACTAATACTCTCGCATTTGTACATAATACACTACGCGTGGGACGA 121807

QY 609 GATAAAAAAGAGACTCAAAATTTCTACTATCCAGAACACTGGATTAACCTCTCTGAGA 668
DB 121808 GAAGACCCAGCACTGATAGTTTTTACTTACCCAAAAGCGCTGCTCTCTCCACTCCGT 121867

QY 669 TTACACAGCCCTGAATCAATCCCTGGGGGCTGATTTTGTATGTCACACAGATTTCAA 728
DB 121868 TTACATANTATAGAGTCCATCTCTACGGAAAGGATTTTAATTTATGCTCAACAGTTTAA 121927

QY 729 CAGCTGTATATGGAGGCTCTGAAAAAGATATCAAGATTTGTGTGACAACTTCCAGGAT 788
DB 121928 ACACCTGGATCTTGAAGCAGTAAAGAAAGATATAAAAAACAGTTCTTACCACATCACAAGAC 121987

QY 789 TGTGTCCCTCGCGATATGTCATATATGTCCTCTTTTATTCTGATGGCTTGGCAGGCT 848
DB 121988 TGTGTCCCTCTGATATGTAATATATGTCCTCAATTTTATTCTGATGGCTGTCATGCT 122047

QY 849 GCGGGAACATACAGGACATATGATGCGCGGGAGGCGCCAGTGTGTGTGACGACGTTTT 908
DB 122048 GCGGGAACCTACCGCATATATATGATGCGCTGGTGGTGGCGGACGCAAGATTT 122107

QY 909 GAACCGCTGAACAGCTGGCGGATATACGTTAATCTGTGATAAAGCCGCTCGATGCTGTGG 968
DB 122108 GAGCCACTCAATAGCTGGCCAGATAACGCCAACCTTGATAAAGCGCGCTCGGCTCTGTGG 122167

QY 969 CCAGTCAAGAAAAAATACGCTCCAGTATTTCTTGGGGACACTGATGGTCTGACTGGT 1028
DB 122168 CCTATAAAGAAAAATATGCGCTAAAATATCTGCGGCGGACCTGTGATGCTTACAGGC 122227

QY 1029 AATGTTGCCCTTCAATCCATGGGATTTTAAACGCTGGGATTTGCTGGCGGAAGAGAT 1088
DB 122228 AATGTCGCGCTCGAATCTATGGGTTTAAACGCTGGGTTTGCAGGAGGCGCTGAGGAT 122287

```

QY 1089 GACTGGGAGTGGACCTGGTATCTACTGGGGCCCTGGACAAAGCCTCTTCCAGATACCCG 1148
DB 122288 GATTGGCAATCCGATCTGGTCTACTGGGGGCGAGGCAACAAGATGCTGTGATACAGG 122347
QY 1149 GATTAACCGGAACTTCAGAAACCTCTTGGCGCCACGAGATGGGACTTATTTATGTC 1208
DB 122348 GATTAACCGGAACTTCAGAAACCTCTTGGCGCCACGAGATGGGACTTATTTATGTC 122407
QY 1209 AATCTGAAGCCCGGTGGAAACACAGATCTCTTGGCTTCCGCGAAGATATACAGGAA 1268
DB 122408 AACCTGAAGCAACCAATATGTAACACAGACCCGGTCTGCGGCAAAAGATATTCGTGA 122467
QY 1269 GCTTTTTCACGTATGGCCATGGATGATCAGAGACTGTGCGCTGTATCGCGGAGGCGAT 1328
DB 122468 CGGTTTTCACGTATGGCCATGGATGATCAGAGACTGTGCGCTGTATCGCGGAGGCGAT 122527
QY 1329 ACATTTGGTAAAGCAGATGCTGCGAGCTCTCTCTGAAATAATGATTTGCGGAGGCGCTGAT 1388
DB 122528 ACCTTCGGTAAAGCAGATGCTGCGAGCTCTCTCTGAAATAATGATTTGCGGAGGCGCTGAT 122587
QY 1389 GGTGACACCTGTGGAGGAGCAGGAGTGGATGGAAATAATGATTTGCGGAGGCGCTGAT 1448
DB 122588 GAAGGCGGCTAGAACACAAAGGTTAGGATGGGCGAATAATGTTGCTCAGGTAAATGTT 122647
QY 1449 AATATACCATCAGGAGTGGCTGGAAGGAGCCTGGTGCAC 1489
DB 122648 AAGGACAGTAAACAGTGGGTGGAAGGCGATGGACAAC 122688

RESULT 10
LOCUS AB017595 2587 bp DNA linear BCT 06-AUG-1999
DEFINITION Legionella pneumophila gene for catalase-peroxidase, complete cds.
ACCESSION AB017595
VERSION AB017595.1 GI:4996127
KEYWORDS kata; catalase-peroxidase.
SOURCE Legionella pneumophila (strain:AM511) DNA.
ORGANISM Legionella pneumophila
Bacteria; Proteobacteria; gamma subdivision; Legionellaceae group; Legionellaceae; Legionella.
REFERENCE 1 (sites)
AUTHORS Anemura-Makawa, J., Mishima-Abe, S., Kura, F., Takahashi, T. and Watanabe, H.
TITLE Identification of a novel periplasmic catalase-peroxidase Kata of Legionella pneumophila
JOURNAL FEMS Microbiol. Lett. 176 (2), 339-344 (1999)
MEDLINE 9356723
REFERENCE 2 (bases 1 to 2587)
AUTHORS Anemura-Makawa, J. and Watanabe, H.
TITLE Direct Submission
JOURNAL Submitted (10-SEP-1998) Junko Anemura-Makawa, National Institute of Infectious Diseases, Department of Bacteriology, Toyama 1-23-1, Sijuku-Ku, Tokyo 162-8640, Japan (E-mail: jmaekawa@nih.go.jp, Tel: 81-3-5285-1111, Fax: 81-3-5285-1163)
FEATURES
source location/Qualifiers
1. 2587
/organism="Legionella pneumophila"
/strain="AM511"
/db_xref="taxon:446"
gene 203..2452
/gene="kata"
CDS 203..2452
/gene="kata"
/codon_start=1
/product="catalase-peroxidase"
/protein_id="BA78342.1"
/db_xref="GI:4996128"
/translation="MKRTIPLEAETLAISPSIFNYAHQEDKPTQYWPVKMLD LSPURQPNATSNPMGEKFNFAEFNSLDLKLMTTSQDWPADYGNYPGLF IRMSHAAGATRYIDSGGANGGFPAPONSFPDNDKRLRLWLP1KQYGRKIS WADLLVLVAGNVAESMGFKTIGFAGGDEAWAENINWGPKEWLSKPDKDKLEK PLAAVTMGLIYNVPEGNVPDPLAAAEKIRETFGRMANMDEETVALIAGHAFGKTH

GAAGKYLGPAPAEAGIEEOGFWKNSYSGSGKDKTITTSLEGAWTVPTRHSHNYLQ
NLEFNWVKTSPPGAAIQWVPENSNASWVDAFDPSKRHPAVMLTDLALKEFDVYS
KIAKFLNPFEEFDDAFARWFLKLRHDMGPRSYLGLSLVPKEAMTQWDPVVDYKL
VDANDIANLKKILNSGLTTSSELVKTAWASASTFRGTDMRGANGARILAPQDWFA
NDPQLAKVLKTESIQNNFNNAQADGKKISLADLIVLGGNAATEQAKAGYDIIIVP
FTPGRLDVTQGMTDVKSPEVLEPRADGRNFTDKSRNNMSPPEMLVEKASLLKLSVPEM
TVLVGGMRVLNANTQONQGVFTDKPGLNNDFFINLLSMSTEMWKSSEETGIEGYE
RRTGKLKWKATSVDLIFGANSSELRAVEAYATDDAKEKFIQDFINAWVKVMTADRFDI
KAANANINS"
BASE COUNT 784 a 551 c 611 g 641 t
ORIGIN
Query Match 25.8%; Score 384.8; DB 1; Length 2587;
Best Local Similarity 65.8%; Pred. No. 2.6e-103;
Matches 592; Conservative 0; Mismatches 30; Indels 6; Gaps 2;
QY 590 TTCTACCGCTGTAGCGCTGATAAAAAAGAGACTCAAAATTTCTACTATCCAGAAACACT 649
DB 271 TTATGCCCATGCGCAAGAGACAGCCGAAACCAATCAATCTGCTGGCCTAAATGCT 330
QY 650 GGATTTTAACCTCTCTGAGATTACACAGCCCTGAATCAATCAATCCTGGGGGCTGATTTGA 709
DB 331 CGATCTTAGCCGCTCGCGCAACCAATGCCACTTCGAATCCATGGGTGAAAAATTCAA 390
QY 710 TTATGCCACAGATTTCAACAGCTGGATGAGGCTCTGAAAAAGATATCAAAAGATTT 769
DB 391 CTATGCCGAAGAATCAATAGCTTAGATTTGAATGCCGTGATGAGATCTCAAAAATTT 450
QY 770 GCTGACAACTTCCAGGATTTGCTGCGCTGCGGATTTATGCTATTTATGTCCTTTTAT 829
DB 451 AATGACTACCTCGAAGACTGGTGGCTGCTGATTTATGTTAACTATGTTCCATTTATCA 510
QY 830 TCGTATGGCTTGGCAGGTCGCGGAACATACAGGACATATGATGCGCGGAGGCGCCAG 889
DB 511 ACGCATGTCTATGCTGCGCAGGAACTTACCGCATCTATGATGCGCGGTGGTGACCAA 570
QY 890 TGGTGGTTCAGCAACAGTTTGAACCGGTGAACAGCTGCGCGGATPAACGTTAATCTGGATA 949
DB 571 CGTGGCTTCCAGGTTTCCGCCCGCAACACAGCTGCGCGGATAATGCCAATCTGGACAA 630
QY 950 AGCCGCTGATTTGCTGGCCAGTCAAGAAAAAATACGCTCCAGTATTTCTCTGGGGAGA 1009
DB 631 AGCCAGACGTTTACTGTGGCCAAATTAAGCAAGAAATATGGGCGCAAGATTCATGCTGCTGA 690
QY 1010 CCTGATGCTCTGATGCTAATTTGTCCTTGAATCATCATGGGATTTAAACGCTGGGAT 1069
DB 691 TTGTTGCTGCTGCTGGAATGTTGCGATGATCTATGGCTTCAAGACCATTTGATTT 750
QY 1070 TGTGCGCGAAGAGAGATGACTGGGAGTGGACCTGGTATATCTGGGGGCTGACACAA 1129
DB 751 TGCTGGAGGCGGTGAAGACGCTGGAGGCTCATCAATATCAATTTGGGGCCCTGAAGGCAA 810
QY 1130 GCCTCTTGCAGATAA---CCGGGATAAAACGGGAACTTCAGAAACCTCTTCCGCGCAC 1186
DB 811 ATGCTGGAAAGTAGAGCCCGAGGATAAGATGGAATACTTGAATAACCGCTTCTGCTGGAC 870
QY 1187 GCAGATGGGACTTATTTATGTCATCTCTGAAGGCCCGCTGGAAACACAGATCTCTGCG 1246
DB 871 CGTATGGCTTAACTATGATGTAATCCGGAAGGACCAACAGCGCTTCCAGATCTCTTTCG 930
QY 1247 TTCGCGGAAAGATATCAGGGAAGCTTTTACGTATGGCCATGATGATGAGGAGACTGT 1306
DB 931 TCGCGCCGAAAAAATTCGAGAGACTTTTGGGCGGTATGGCCATGAATGACGAAGAACAGT 990
QY 1307 GGCCTGTATCGCGGAGGCGCATACATTTGTTGAAGACATGCTGTCAGCGCTCTCTGAAA 1366
DB 991 TGTCTTAAATGCGGCGGCGCAGCATTCGGAACACACATGAGCGCGCT---CCGCA 1047
QY 1367 ATGTATTGCGCGAGGCTGATGTTGTCACCTGTGTGAGGAGGAGGAGGACTGGATGAAAA 1426
DB 1048 ATATCTGGGCGCGCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1107
QY 1427 TAAATGCTGTACAGGAACCGCAATATATACCATCAGGAGTGGCTTGGGAGGAGGAGGCTG 1486

```

Db      1108 CAGTTACGGCAGCGGCAAGGAAAGACACAGATTACCACTGGTTAGAGAGTGCCTGGAC 1167
|||||
RESULT 11
AF276752
LOCUS
DEFINITION
Legionella pneumophila catalase-peroxidase (kata) gene, complete
cfs.
ACCESSION
AF276752
VERSION
AF276752.1 GI:11528084
SOURCE
Legionella pneumophila.
ORGANISM
Legionella pneumophila.
Bacteria; Proteobacteria; gamma subdivision; Legionellaceae group;
Legionellaceae; Legionella.
REFERENCE
1 (bases 1 to 5249)
Bandyopadhyay, P. and Steinman, H.M.
Catalase-peroxidases of Legionella pneumophila: cloning of the kata
gene and studies of KATA function
J. Bacteriol. 182 (23), 6679-6686 (2000)
JOURNAL
MEDLINE
20528318
PUBMED
11073912
REFERENCE
2 (bases 1 to 5249)
Bandyopadhyay, P. and Steinman, H.M.
Direct Submission
Submitted (09-JUN-2000) Biochemistry, Albert Einstein College of
Medicine, 1300 Morris Park Avenue, Bronx, NY 10461-1602, USA
LOCATION/Qualifiers
1..5249
/organism="Legionella pneumophila"
/strain="JR32"
/db_xref="taxon:446"
misc_difference 2225
/notes="compared to Legionella pneumophila kata sequence in
GenBank Accession Number AB017595"
/replaces="g"
misc_difference 2227
/notes="compared to Legionella pneumophila kata sequence in
GenBank Accession Number AB017595"
/replaces="c"
gene
2386..4635
/gene="kata"
CDS
2386..4635
/gene="kata"
/notes="kata; localized in periplasm"
/codon_start=1
/transl_table=11
/product="catalase-peroxidase"
/protein_id="AAG37106.1"
/db_xref="GI:11528085"
/translation="MFKRTIPLFAFTLAISIFPNVAHAQEDKPKTNQYWPVKMLD
LSPLQPNATSNPMGEKNYAEFNLSLDNAVIEDLKLMTTSQDWPPADYGNYPFLF
IRMSWHAAGTYRIYDGRGANGGQFORFAPSNPDNANLDKARLLPWLKQYGRKIS
WADLLVLAGNAMESMGPKTIGFAGREDWEA INNWGPEKWLKESKRODKGLEK
PLAATVMGLIYVNPSPGNCVDPDLAAAEKIRTFEGMANNDEETVALLAGHAFKTH
GAASRYLGPAPAEAGIEQGEQWNSYSGSGKDTITSGLEGATVTPHWSHNLQ
NLFNWNVKTSPGGAQVNPENSNASSVPDAPFSPKHPHVMLTLDALKEDPVYS
KLAFLRDNKPEFDLAFARAFKLIHRDMGPRSYGLSLVPKEMAMTQDPVPVYKL
VDANDIANLKLNSGLTSELVKTANASASTFRGDMRGANGARIAPQKDWPA
NPQELAKVLKTLSTONNFNAQDGKKISLADLVLGGNAAIEQAQKQCYDLIVP
FTPGRTDQVNTGDKVSEVLEPKADGPNFYDKSNMSPPEMKLEKASLLKLVPEM
TVLVGDMRVLNATGQNVQFDPKPTGLNNDFFINLMSWTEWKKSSSTEGIYEGE
RTGKLKWKATSVDLIFGANSSELVAEAYATDDAKERFIQDFINAWKVMYADRFDI
KAANAIN"
BASE COUNT 1771 a 1058 c 1095 g 1325 t
ORIGIN
Query Match 25.8%; Score 384.8; DB 1; Length 5249;
Best Local Similarity 65.8%; Pred. No. 3.1e-103;
Matches 592; Conservative 0; Mismatches 302; Indels 6; Gaps 2;
QY 590 TTCTACCGCTGTAGCCGCTGATATAAAAAAGAGACTCAAAATTTCTACTATCCAGAAACACT 649

```

```

Db      2454 TTATGCCCATGGCAAGAAAGACCGGAAAAACAATCAATACTGTGTGCCTAAAAATGCT 2513
|||||
QY      650 GGATTTAACTCTCTGAGATTACACAGCCCTGAATCAAAATCCCTGGGGGGCTGATTGTGA 709
|||||
Db      2514 CGATCTTACCCCGCTCGCGCAACCAATGCCACTTCGATCCATCCATGGGTGAAAAATTCAA 2573
|||||
QY      710 TTATGCCCAACAGATTCAACAGCTGGATATGGAGGCTCTGAAAAAGATATCAAGATT 769
|||||
Db      2574 CTATGCCGAAGAATTCATAGAGTTAGATTGAATGCGGTGATAGAAGATCTCAAAAAATT 2633
|||||
QY      770 GCTGACAACTCCACGAGATTGTGCGCTCGGATTATGTCATATTGGTCTCTTTCTTAT 829
|||||
Db      2634 AATGACTACTCGAAGACTGGTGGCTCTGATTATGTGTAACTATGGTCCATATTTCAT 2693
|||||
QY      830 TCGTATGGCTTGGCACGTCGCGGAACATACAGGACATATGATGGCGGGGAGCGCCAG 889
|||||
Db      2694 ACGCATGTCTGCGATCGCGCAGGAACATACCGCATCTATATGGCGGTGGAGCAAA 2753
|||||
QY      890 TGGTGTCTAGCAACGTTTTGAACCGCTGAACAGCTGGCGGATAACGTTAAATCTGATAA 949
|||||
Db      2754 CGGTGGCTTCCAGCGTTTCGCCCGCAAAACAGCTGGCGGATAATGCCAATCTGGACAA 2813
|||||
QY      950 AGCCGCTCGATTGCTGCGCCAGCTCAAGAAAAAATACGGCTCAGTATTCTCTGGGAGA 1009
|||||
Db      2814 AGCAGACCTTTACTGTGCCAATTAAGCAGAAATATGGCGCAAGATTTTCATGGCTGA 2873
|||||
QY      1010 CCGTATGGTCTGACTGGTAAATGTTGCCCTTGAATCCATGGGATTTAAACCGCTGGGATT 1069
|||||
Db      2874 TTTGTTGGTACTTGTCTGGAATGTTGCGATGGAATCTATGGGCTTCAAGACCATTTGATT 2933
|||||
QY      1070 TCGTGGCGGAAGAAGATGACTGGGAGTCGGACCTGATATCTACTGGGGCTGACAAACA 1129
|||||
Db      2934 TCGTGGAGGCGCTGAAGACGTCATGGAGGCCATCAATATCAATTTGGGGCTCTGAAGGCA 2993
|||||
QY      1130 GCCTCTTCGACAGATAA---CCGGGATAAAACGGGAAACTTCAGAAACCTCTTCCGCCAC 1186
|||||
Db      2994 ATGGCTGGAAAGTAACGCCCGCAGGATAAAGATGAAAAAACCCTTGAACACCGCTCTCGAC 3053
|||||
QY      1187 GCAGATGGGACTTATTATGTCAATCCTGAAGGCCCGCGGTGGAAAAACACAGATCCTCTGGC 1246
|||||
Db      3054 CGTAATGGGCTTAATCTATGTGAATCCGGAAGGACCAACACGGGTTCCAGATCCTCTTGC 3113
|||||
QY      1247 TTCGGGAAAGATATCAGGGAAGCTTTTTCACATATGCGCCATGGATGATGAGAGACTGT 1306
|||||
Db      3114 TCGGCGCAAAAAAATTCGCGAGACTTTTCGGGCGTATGGCCATGAATGACGAGAAACAGT 3173
|||||
QY      1307 GGCCTGTATCGGGGAGGCATACATTTGTAAGACACATGTCGACGCTCTCTGAAAA 1366
|||||
Db      3174 TCGTTTAATTCGGGCGGCGCACGATTCGAAAAAACACATGAGCGCGCT---CCGGCAA 3230
|||||
QY      1367 ATGTATTGGCGCAGGCGCTGATGGTCACCTGTGAGGAGCAGGACTGGGATGGAAAAA 1426
|||||
Db      3231 ATATCTGGGCGCAGCGCCAGACGACGACGACATAGAAGAACAGGCTTTGGTTGGAAAAA 3290
|||||
QY      1427 TAAATGTGTACAGAAACCGCAATATACCATCACAGTGGCCTCGAAGAGCCTGGTTC 1486
|||||
Db      3291 CAGTTACGCGCGGCAAGGAAAGACACGATACCACTGGTGTAGAAAGTGCCTGGAC 3350
|||||
RESULT 12
ECIS91TP/c
LOCUS
DEFINITION
Escherichia coli plasmid pSU2911 including tnpa gene for
transposase and ORF121.
ACCESSION
X17114.5 GI:21038833
VERSION
X17114.5
KEYWORDS
insertion sequence; tnpa gene; transposase.
SOURCE
Escherichia coli.
ORGANISM
Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
Escherichia.
REFERENCE
1 (bases 1 to 1829)

```



```

/db_xref="GI:19914794"
/translation="MOOSLSHSLFPLCGRAFORDDLMERNTCMILALDVTDRREALKI
AEDVWEFVAIKVYPLILATIGLIIRELAEPFVIADKFADIPNTRNLICDQFEA
GADAVLAQPTGDRSDACIEIASEYKDFVYVSEMSHPGGADFPQVAEAKWALE
AGAGLVAPOPTGRKVRKRIIRIIGDKLTIIISPCVGAQGSADVIAAGADWIVORSI
YKASPEKARQIAEILQELRGEY"
5307..7799
/gene="MA0970"
CDS
5307..7799
/gene="MA0970"
/codon_start=1
/transl_table=11
/product="sensory transduction histidine kinase"
/protein_id="AA04402.1"
/db_xref="GI:19914795"
/translation="MOEYSKINSQKGIQKGMNAKVQTPATNOLVLSVDNNGTPI
IHSNEASEPLTKMGVSGVEKLPSSIVDLQWVSRNSPEKTEVGNKLYLVFSPIL
PEQRYVISGFDSDOKPEKMGVSETRKARDLVEELRSEKFTLFTIDEGF
CVTEVLEDNDPCIDYRFLIENNSFERQGIKNVGRMRREIAPEHEEHWFQIYNGVA
LTGEPVRFENPALGHHFYDVAFRTGDDPHRRVGILFNDIAKRKRAEEMLESARQ
AFLLQSDALRPLFDIEIOAVATRVILGHELJARVAYGEVDDDTFIVFERNYVAPG
VPLVTRFRMADPGLLSALKKEGRTIVVPDLASSTELSEASERSYANIGIASLMGVP
LVKGGRFVLAHVHTAPQWSPAYVALIETEARTWAVERAKAEALKESEERFTL
LSTSPGVGVSSAESKIIYTPAYDRILGNPESELIGNSDTVNPERKSWRCTL
KRGVVRDETFEKKKEDKTPVWVSINTSPISFGGRKAAMGTIQDISERQAREALRES
ELKFRILSNLEKVERVELEKAYKFLQEIIRKQEIHRKRNKQVLSLDDLOQ
AEFRKKNEDSKILEAFKESQDRVISMSALTHEELHKSGEIDTLNFSAYIHELSGN
FLSYRNGDISIDMDIEDIFDDMTSVPLMIVNELVNSLKNHAFSDRDKEIRIK
LHRKNRESIDECCLAFILSDNGIGIPKDLIEIDIESLGLQLVLTLDLQDELE
LKDDGTEFTVRESITERNQANSPTQQNEF"
complement(8223..8672)
/gene="fur"
/note="MA0971"
complement(8223..8672)
/gene="fur"
/codon_start=1
/transl_table=11
/product="ferric uptake regulation protein"
/protein_id="AA04403.1"
/db_xref="GI:19914796"
/translation="MSNLQSDPDLIIKALRGKYKATPQRIAGIOFVLNHAHPTAQR
IYSEVKVYPTVSLATYTKTVQLKEVGLIQFNLEKQDQRPNEPNAHLVCLQCK
SINDCTDPMISEVDVDRMSNEVDVFSAGEWNFNIFGICSNCRKKSICD"
8913..11123
/gene="MA0972"
CDS
8913..11123
/gene="MA0972"
/codon_start=1
/transl_table=11
/product="catalase/peroxidase"
/protein_id="AA04404.1"
/db_xref="GI:19914797"
/translation="MGNYMTDDKMSVTSKANKQETGRDMSNRDWWPHLKLILHQ
HSSKSNPMDENFYAKFEKSLDLAAVKDLAALMDSQDWPADFGHYGLPIRMAWH
SACTYRAGDGGGGGQOQREAPLNSPDNVLNDRALLWPKOKYGRKISWADMI
LTCNVAMETWFKTFGGGREDYWEPDQDYWGSEDTWLGDERITGDRDLNPLAAY
QMGLIIVPEGNPDPIAAKNDIREVFARMAMNDEETVALIAGHAPKFKTHGAPPA
SHVGPPEEAIEAQGLGWKSSFGTGKDDTITGGLEVTVNTPTKWSNFFRILFGY
EWELTSPAGAYOWKPKGGAGACTIPDAHDPKSRHAPSMNTDLSLRFDPVYKISRH
FYENPQLDASFAWFWFKLTHRMDGPRARYLGPVPAELIWDQPTPAVNHLEIDKQ
IATFKRILASGLSISOLYSTAWVSASTFSGDKRGGANGARIRLAPOKDWNVPAPAE
LAKVNLTLGIESIEFNSAASGGKKVSLADLIVLACGAGVEQAQKNGVDVTPFLFGR
MDALQCTDVVSFALLEPAGDFRNLYLKAQYPLPAEELLVDRKAQLTLTAPETVLVG
GMVNTVHTGHQVFTQKPEALTNDFVNLMDTEWKAYSVDKIDPEGCDRKTGE
VKWTGRVDLIFGNSQLRALAEVYGSADAQEKVQDFVAAATKVMNLDREFDLA"
BASE COUNT 3077 a 2495 c 2699 g 3049 t
ORIGIN

```

Query Match 23.5%; Score 349.4; DB 1; Length 11320;
 Best Local Similarity 66.6%; Pred. No. 1.5e-92;
 Matches 548; Conservative 0; Mismatches 266; Indels 9; Gaps 3;
 672 CACAGCCCTGAATCAATCCCTGGGGGCTGATTTTGTATTATGCCACCAGATTCAACAG 731

```

Db 9045 CATTCTCTCAAGTCCAAACCCGATGGCGGAGGACTTCAATTACGGAAGGACTTCAAGAGC 9104
Qy 732 CTGGATATGGAGGCTCTCAAAAAGATATCAAAGATTTGCTGACAACCTTCCCGAGATTGG 791
Db 9105 CTAGATCTGGCGGCTGTGAAAAAGATCTTCGGCGGCTGTAGACCGATTCGACGAGCTGG 9164
Qy 792 TGCCCTGCGGATTATGTCATTATGTGCTCTTTCTTTTATTCGTATGGCTTGGCAGCGTGCC 851
Db 9165 TGGCGCGGGATTTTGGCCATTATATGGCCCTTTGTTTATTCGATGGCATGSCACAGCGCA 9224
Qy 852 GGAACATACAGACATATGATGGCGGGGAGCGCCAGCTGGTGTGTCAGCAACGTTTGA 911
Db 9225 GGCACCTTATCGCGCGGTGACGGCGCGCGCGGTAGAGGCCAGCAACGCTTCGCG 9284
Qy 912 CGCGTGAACAGCTGCGCGGATTAACCTTAATCTGGATAAAGCCGTCGATTCTGTGGCCA 971
Db 9285 CGGCTCACAGTGTGGCGGACANVTCAACTGACAAAGCGGCTGCTGCTCTGTGGCG 9344
Qy 972 GTCAAGAAAAAATAGGCTCCAGTATTTCTTGGGAGACCTGTATGGTCTGCTGACTGGTAAT 1031
Db 9345 ATCAAGCAGAAGTATGGCGGAAAAATTTTCATGGCGCGACCTTATGATTTCTTACGGCAAC 9404
Qy 1032 GTTGCCTTGAATCCATCGGATTTAAACGCTGGGATTTGCTGGCGGAGAGAAGATGAC 1091
Db 9405 GTCGCAATGGAAACTATGGGATTCAAAACGTTTGTGTTTCGGGGGCGGTCGCGAGGACGTC 9464
Qy 1092 TGGGAGTCGGACCTCT--GTATCTATGGGGGCTGACAAACAGCCCTCTTTCAGATAAACCGG 1148
Db 9465 TGGGAGCGGACCGAGGACGCTATTTGGGGCTCTGAGGACACGTGCTGGGTGATGACGC 9524
Qy 1149 GATA---AAAACGGGAAACTTCAGAACTCTTGGCGCCACGACGATGGGACTTATTTAT 1205
Db 9525 TACACCGCGGACCGGGACCTCGAAAATCCGCTCGCTGCGGTGCGAGTGGGCTGATCTAT 9584
Qy 1206 GTCAATCTGAAGGCCCGCGTGGAAACACAGACTCTCTGGCTTCGCGGAAAGATATCAGG 1265
Db 9585 GTCAATCTGGAGGCCCGGACGACGACGATCGGATCGATAGCGGCGGCGCAAGATATCGCG 9644
Qy 1266 GAAGCTTTTTCAGTATGCCATGGATGATGAGGAGACTGTGGCCCTGATCGCGGAGGG 1325
Db 9645 GAGGTTTTTCGCTCGCATGCCATGAACACGACGAGAGACCGTTGCGCTCATCGCTGGCGGT 9704
Qy 1326 CATACATTTGTAAGACACATGGTGACCGCTCTCTGAAAAATGATTTGGCGGAGGGCCT 1385
Db 9705 CACGCTTCGGAACAAACCCACGCGTCCG---GCCCTGCGTCCCATGTGCGGCTCGAGCA 9761
Qy 1386 GATGTGCACCTGTGGAGGACGAGGACTGGGATGGGATGGAATAAATAATGTGTACAGGAAC 1445
Db 9762 GAAGCGGCGCATGTCAGCGCGCAGGCTCGGTTGGAGACAGCAGCTTCGCGACTGGCNA 9821
Qy 1446 GGCAATATACCATCACCAGTGGCTTGGAGGAGCGCTGGTCTGA 1488
Db 9822 GSGCATGACAGCATCACCGCGGCGCTGGAGGTGACCTGACCTGGACCA 9864

```

RESULT 15
 AF294823
 LOCUS
 DEFINITION
 Shigella sonnei O antigen gene cluster, complete sequence;
 insertion sequences IS1 InsB (insB), InsA (insA), IS630
 transposase (insA), and IS911 InsB (insB) genes, complete cds.
 AF294823
 AF294823.1 GI:15149169
 SOURCE
 Shigella sonnei.
 ORGANISM
 Shigella sonnei
 Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 Shigella.
 1 (bases 1 to 17986)
 Xu,D.Q., Cisar,J.O., Ambulos,N. Jr., Burr,D.H. and Kopecko,D.J.
 Molecular Cloning and Characterization of Genes for Shigella sonnei

Form I O Polysaccharide: Proposed Biosynthetic Pathway and Stable Expression in a Live Salmonella Vaccine Vector
Infect. Immun. 70 (8), 4414-4423 (2002)

121117952

2 (bases 1 to 17986)

Xu,D.Q., Cisar,J.O., Ambulos,N. Jr., Burr,D. and Kopecko,D.J.
Direct Submission
Submitted (09-AUG-2000) Oral Infection and Immunity Branch,
National Institutes of Health, National Institute of Dental and
Craniofacial Research, Bldg. 30, Room 302, Bethesda, MD 20892, USA

JOURNAL
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL

FEATURES
source
1..17986
/organism="Shigella sonnei"
/strain="53G"
/db_xref="taxon:624"
/clone="cosmid pWR101"
1..768
/note="similar to Escherichia coli IS1"
/insertion_seq="IS1"
complement(16..519)
/gene="insB"
complement(16..519)
/gene="insB"
/note="similar to Escherichia coli IS1 InsB"
/codon_start=1
/transl_table=11
/product="InsB"
/protein_id="AAK85161.1"
/db_xref="GI:15149170"
/translation="MFGNCTHYGRWQHDFTSLKKLRPOSVTSRIQPSGSDVIVCAEMD
EQWYVGAKSRQRLVAYDSLRTYVAHVFGERTWATLGRMLSLSPDFVIVMWDG
WPLYESRLGKGLHIVISQRTQRTQRTIRNLRLQHLARLGRKLSFSKSKVELHDVKVIGHY
LNKHYQ"
complement(438..713)
/gene="insa"
complement(438..713)
/gene="insa"
/note="similar to Escherichia coli IS1 Insa"
/codon_start=1
/transl_table=11
/product="Insa"
/protein_id="AAK85162.1"
/db_xref="GI:15149171"
/translation="MASVTSICPSCSATDGVWNGKSTAGHQRYLCSHCRTKWLOFT
YTAQPCTHQKIIDMAWNGVCRTARIMGVGLNTILRLKNSGRS"
complement(721..724)
complement(734..739)
complement(756..761)
788..1720
/gene="wzz"
788..1720
/gene="wzz"
/note="similar to Shigella sonnei rfbR and Actinobacillus
actinomycetemcomitans wzz proteins"
/codon_start=1
/transl_table=11
/product="truncated wzz protein"
/protein_id="AAK85163.1"
/db_xref="GI:15149172"
/translation="MFAFTAKEQWTKTIQAPRIAEGLSYLFKHQAYARILNOPLDT
NALANGFLSDILIIAQSPTKVFLESTYTKRETNNLTDQDKKTLWLAQNKGLVI
TPKEKTSYVIITQASDAQAYKLLQYLKANNVNAVTLSLDEFGQNVTLPLGK
NKEI IDTFQRSEKLIDIAHQIORDTLTFAQGI IDVRSSKGGFDNAQSSYKPLIGEK
LLSAELKATKDAPIIYFPFYEVKROIDEGLMRDNIQAQAYRYQMKPSEPVIKDKP
NKAILLIGALPAMFAIVGLTVALTKDKTKLD"
1645..1650
1666..1671
/note="putative promoter of O-antigen biosynthetic operon"
1745..3069
/gene="wbgt"
1745..1749
/gene="wbgt"
1756..3069

-35_signal
-10_signal

gene
RBS
CDS

```

RBS      6885.. 6887
CDS      6894.. 7925
        /note="ORF8; Tpsae; smilar to ts630 53g ORF5 and HW383
        /transposase"
        /codon_start=1
        /transl_table=11
        /product="transposase"
        /protein_id="AAK85168.1"
        /db_xref="GI:15149177"
        /translation="MPIIAPISDRERLMQKAIHKTHDKNYARRLTAMLMLHRGDRVS
        DVARTLCARSSVGRVWNFTQSGVEGLKSLPAGRARWPFEEHICTLLRELKHSFGD
        FGVRQSSVELLAIKINETGCOLNASTVRWLPAGIWRAPPTLRDPHDKDEK
        MAIHKALDCSAREHVFYDEVDIHLNFKIGDWQLRQWRARPTPGQNEKAYILAGA
        LHSGTGVCVGNSSKSLFISLLRLKATYRRAKTIITLVDNYIIHKSRSTQSWLK
        ENPKFRVYQVTPVSPMNHVRELWQALHDTITRNHOCSSMWQLLKVRHFMETVSPFP
        GGRHGLAKV"
        /length=1202
gene      7947.. 9202
        /gene="wbgv"
RBS      7947.. 7950
        /gene="wbgv"
CDS      7938.. 9202
        /gene="wbgv"
        /note="ORF9"
        /codon_start=1
        /transl_table=11
        /product="wbgv"
        /protein_id="AAK85169.1"
        /db_xref="GI:15149178"
        /translation="MLLEYVERKISIALSKYPRVDVIRFFVLYIASLPGIILNKNT
        VIOQKIVEISIDSESEFFCYDHPMSNGRVLFPSSAPFKRHPKVKYISICVK
        DLLNKKVLYIDPRAFNWQOGRSLMWDIDDDNIIFENDYENNGIISVYSLSMKVIKKI
        NYPIDVNNYKATLDFSLAKYDSYGYNNKSFSDISINLNTGGIELFSLDDEM
        LKTNFKCNIDVHVNHFMAFDPGRSVMFIHRYTPKGRERLHWNLINDNVRVLI
        NESIISHCWGNDEIIFGFAEISLNTYRILSIECNTEKLFDFARKYSQGHPTIVH
        NRYIISDTPDKNRKIKFLFYDLVKNDRYRELGLFYESMFFSYSCDLHPRTISVDNRE
        LFYDSVHSGRKKRKYFMRSGICE"
        /length=10181
gene      9181.. 10181
RBS      /gene="wbgv"
        9181.. 9184

Query Match      23.4%; Score 348.8; DB 1; Length 17986;
Best Local Similarity 93.1%; Pred. No. 2.5e-92;
Matches 376; Conservative 0; Mismatches 27; Indels 1; Gaps 1;

QY      4 CAGTCGGGAGATGAA-AGCACCACTGTGTATACCCCATCAGCGTGTCTCCGAGGCCAGT 62
Db      16534 CAGTCGGGTGATGAACAGCACCACTGTGTATACCCCATCAGTGTGTCTCCGAGGCCAGT 16593

QY      63 ATTTTGTGCACACTCAATGACTACCGGACGACATGAACCTTCGGGTGTTTCTCCAGC 122
Db      16594 ATTTTGTGCACACTCAATGACTACCGGACGACTGCCCCCTTCGGGTGCTTCTCCAGC 16653

QY      123 CAGTTAAGCCAGCGGTTTCCCTGCTGAAATAATGTCGGCAAAACGGGGAAGCATCAGAAGG 182
Db      16654 CAGTTAAGCCAGCGGTTTCCCTGCTGAAAGATATCGCAAAACGGGGAAGCATCAGAAGG 16713

QY      183 CGCGGGGAACCTCGTCGGCAGTGAACCGTGCACACTCCGGGAGTACATACGCGCGG 242
Db      16714 CGCGGGGCACTCGTCGGCAGTGAACCGTGCACACTCCGGGAGTACATACGCGCGG 16773

QY      243 CGCTGATACCGGCAAGAATGTTGCGCAAACTCCCGCTCCGTCCGAGCGGCTATTTCAGGAT 302
Db      16774 CGCTGATACCGGCAAGAATGTTGCGCAAACTCCCGCTCCGTCCGAGCGGCTATTTCAGGAT 16833

QY      303 ACCCTTCGTATCAACACGCTACAAACGAGAAGACCAAGCTTTTGTGTTCTGATCCACAA 362
Db      16834 ACCCTTCGTATCAACACGCTACAAACGAGAAGACCAAGCTTTTGTGTTCTGATCCACAA 16893

QY      363 AGAAGGATATTCAGGCTGCGCAGCACTCAACGCGCATCGTCA 406
Db      16894 AGAAGGATATTCAGGCTGCGCAGCATTCACGCGCATTCGATCAACGCGCATCGTCA 16937

```

```

RESULT 16
LOCUS   AP001510/c
DEFINITION
Bacillus halodurans genomic DNA, section 4/14.
AP001510 BA000004
VERSION AP001510.1 GI:10173440
KEYWORDS
SOURCE  Bacillus halodurans DNA.
ORGANISM
Bacillus halodurans
Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
REFERENCE
1 (sites)
AUTHORS Takami,H., Nakasone,K., Hiramata,C., Takaki,Y., Masui,N., Fuji,F.,
Nakamura,Y. and Inoue,A.
TITLE An improved physical and genetic map of the genome of alkaliphilic
Bacillus sp. C-125
JOURNAL Extremophiles 3 (1), 21-28 (1999)
MEDLINE 99184645
PUBMED 10086841
REFERENCE
2 (sites)
AUTHORS Takami,H. and Horikoshi,K.
TITLE Reidentification of facultatively alkaliphilic Bacillus sp. C-125
JOURNAL Biosci. Biotechnol. Biochem. 63, 943-945 (1999)
REFERENCE
3 (sites)
AUTHORS Takami,H., Nakasone,K., Ogasawara,N., Hiramata,C., Nakamura,Y.,
Masui,N., Fuji,F., Takaki,Y., Inoue,A. and Horikoshi,K.
TITLE Sequencing of three lambda clones from the genome of alkaliphilic
Bacillus sp. strain C-125
JOURNAL Extremophiles 3 (1), 29-34 (1999)
MEDLINE 99184646
PUBMED 10086842
REFERENCE
4 (sites)
AUTHORS Takami,H., Takaki,Y., Nakasone,K., Hiramata,C., Inoue,A. and
Horikoshi,K.
TITLE Sequence analysis of a 32-kb region including the major ribosomal
protein gene clusters from alkaliphilic Bacillus sp. strain C-125
JOURNAL Biosci. Biotechnol. Biochem. 63 (2), 452-455 (1999)
MEDLINE 99209008
PUBMED 10192928
REFERENCE
5 (sites)
AUTHORS Takami,H., Takaki,Y., Nakasone,K., Sakiyama,T., Maeno,G.,
Sasaki,R., Hiramata,C., Fuji,F. and Masui,N.
TITLE Genetic analysis of the chromosome of alkaliphilic Bacillus
halodurans C-125
JOURNAL Extremophiles 3 (3), 227-233 (1999)
MEDLINE 99411980
PUBMED 10484179
REFERENCE
6 (sites)
AUTHORS Takami,H.
TITLE Genome analysis of facultatively alkaliphilic Bacillus halodurans
C-125
JOURNAL (in) Extremophiles in deep-sea environments (Ed.);
: HORIKOSHI, K. TSUJII;
: 249-284; Springer-Verlag (1999)
REFERENCE
7 (sites)
AUTHORS Takami,H., Masui,N., Nakasone,K. and Horikoshi,K.
TITLE Replication origin region of the chromosome of alkaliphilic
Bacillus halodurans C-125
JOURNAL Biosci. Biotechnol. Biochem. 63 (6), 1134-1137 (1999)
MEDLINE 99356711
PUBMED 10427704
REFERENCE
8 (sites)
AUTHORS Takami,H. and Horikoshi,K.
TITLE Analysis of the genome of an alkaliphilic Bacillus strain from an
industrial point of view
JOURNAL Extremophiles 4 (2), 99-108 (2000)
MEDLINE 20263314
PUBMED 10805564
REFERENCE
9 (sites)
AUTHORS Nakasone,K., Masui,N., Takaki,Y., Sasaki,R., Maeno,G., Sakiyama,T.,
Hiramata,C., Fuji,F. and Takami,H.
TITLE Characterization and comparative study of the rrm operons of
alkaliphilic Bacillus halodurans C-125

```


QY	801	GATTATGTCATTATGTCCTTCTTTATTCGATGCTTGGCACGTCGCCGAACATAC	860
Db	83663	GACITTTGGTCATTATGTCCTTCTTTATTCGATGCTTGGCACGTCGCCGAACATAC	83604
QY	861	AGGACATATGATGCGCGGAGGCGCGAGTGGTGGTCAGCAACGTTTGAACCGCTGAAC	920
Db	83603	CGTATAGTGTATGCGCGGAGGTCGTGGCGGAACGTGTAACCAACGCTTTGCCCTCTTAAT	83544
QY	921	AGCTGGCGGATACGTTAATCTGGATAAAGCCGTCGATGCTGTGGCCAGTCAAGAA	980
Db	83543	AGCTGGCCCTGACACGCGCAACCTTGATAAGCCGCGGATCTACTATGGCCAAATAACAA	83484
QY	981	AAATACGCTCCACATTTCTCTGGGAGACCTGATGCTCTGACTGTGTAATGTTGCCCTT	1040
Db	83483	AAATACGCGACAGATTTCTTGGCTGACTTACTTGTATAGCAGTATGAGCCATT	83424
QY	1041	GAATCCATGGGATTTAAACGCTGGGATTTGCTGGCGGAGAGAGAAGTACTGGGAGTCG	1100
Db	83423	GAAGATATGGTGGACCGGCTGATTTGGCTTCGGAGCTGGAGCGGAAGACATTTGGCACCCCT	83364
QY	1101	GACCTGG---TATACTGGGGCCCTGACAAAGCCCTCTTCAGATACCGGGGATAAAAC	1157
Db	83363	GAAGAGGATATTACTGGGGAGCGAAAGAGTGGTTAACGGGTGATAACCTTACTCA	83304
QY	1158	GGGA-----AATTCAGAAACCTCTTGGCCGCCACGACAGATGGACATTTATGTCAT	1211
Db	83303	GGTATCTGTAGCTGTGAGATCCACTTGCAGCTGTTGAGATGGGGCTTATTTACGTTAAC	83244
QY	1212	CCTGAAGCCCGCTGGAAACACAGATCCTCTGCTTCGCGGAAAGATATCAGGAGCT	1271
Db	83243	CCTGAAGTCCAGACGGTAAACACAGATCCTTATAAAGCCGCTCATGACATTCGGGAGACA	83184
QY	1272	TTTTCACGTATGGCATGTGATGAGGAGACTGTGGCCCTGATCGCGGAGGCGCATACA	1331
Db	83183	TTTGGCCGATGGGAATGAAGATGAAGAACGGTTCGCCCTTATCGCTGTGGACATACA	83124
QY	1332	TTTGGTAAAGCACATGTGTGACGCTCTCTCGAAATATGATTTGGCGCAGGCGCTGATGGT	1391
Db	83123	TTTGGTAAAGCTCAGCTGAGGAGAAATCCAGATCA---TGTGTGTCAGAGCGCAAGAGCT	83067
QY	1392	GCACCTGTGGAGGAGCGGACTGGGATGGAAATAAATATGTTGTCAGGAAACGGCAAA	1451
Db	83066	GCTCGGATTAAGCACACAGGCTTAGCTGGCAAAACATACGCTTCGCCAAGGTCGC	83007
QY	1452	TATACCATCACAGTGGCGCTGGAAAGGAGCGCTTGTC	1486
Db	83006	GACACGATCATAGTGGCCCTTGAAGGCGCTTGGAC	82972
RESULT 17			
AF486647			
LOCUS			
DEFINITION	AF486647	3289 bp	DNA linear BCT 25-MAR-2002
ACCESSION	AF486647		Rhizobium etli plasmid pCFN42f katG regulator (oxyR) and katG
VERSION	AF486647.1		(katG) genes, complete cds.
KEYWORDS			complement(6, .947)
SOURCE			complement(6, .932)
ORGANISM			Rhizobium etli.
REFERENCE			Rhizobium etli.
AUTHORS			Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
TITLE			Rhizobiaceae; Rhizobium.
JOURNAL			1 (bases 1 to 3289)
REFERENCE			Vargas,C., Encarnacion,S., Davalos,A., Reyes,A., Mora,Y.,
AUTHORS			Garcia,A., Brom,S. and Mora,J.
TITLE			Only one catalase, katG, has been detected in Rhizobium etli, this
JOURNAL			gene and the regulator OxyR are encoded by a plasmid replicon
REFERENCE			unpublished
AUTHORS			2 (bases 1 to 3289)
TITLE			Vargas,C., Encarnacion,S., Davalos,A., Reyes,A., Mora,Y.,
JOURNAL			Garcia,A., Brom,S. and Mora,J.
REFERENCE			Submitted (22-FEB-2002) Metabolic Engineering, CIPN-UNAM, Avenida
AUTHORS			Universidad s/n, Cuernavaca, Morelos 6210, Mexico

FEATURES			
source			
Location/Qualifiers			
1..3289			
/organism="Rhizobium etli"			
/db_xref="taxon:29449"			
/plasmid="pCFN42f"			
complement(6, .947)			
/gene="oxyR"			
complement(6, .932)			
/gene="oxyR"			
/codon_start=1			
/transl_table=11			
/product="katG regulator"			
/protein_id="AAL93240.1"			
/db_xref="GI:19702288"			
/translation="NKNLTAKLRYFEALRGDFFRAAGSCAVSOPALSMOIKLEQ			
ELGSELPERSREVKLTPFGQTFALRVRDILPAVDELGEFARSDSLFRLRGIIP			
LTAPILLPLVHRLNKSAGIQIEVRETHIKLVHELQAGQMDMAIVAPLVSEAGPTE			
LSLFNRIAGPAPRRKRPVREARLREMLRLLEEGHCFRQDALSPCKVGAARPR			
IMEGSLSTLVQMVAGIGITLIPEMAVPVETTRAHVSIISRPSPAPRTTIGMWLNF			
HPMGNHLPWKLAKARSLVWRKTSPSAINGG"			
complement(942, .947)			
/gene="oxyR"			
1064..3260			
/gene="katG"			
1064..1070			
/gene="katG"			
1077..3260			
/gene="katG"			
/codon_start=1			
/transl_table=11			
/product="katG"			
/protein_id="AAL93241.1"			
/db_xref="GI:19702289"			
/translation="MDNPTDTAGKCPVAHGNKPRGNSNRDWNPNQVLIILHNSGRA			
DPLGKDFDAEDIOEARSRAEKLHALMDSOWPADFGHYGLFIRMAHWSAGTY			
RUTDGGGACGHQHRFAPLNSHPDNALDKAPILLWPKQKGNISHLADILLGNV			
ALDSMGFTGTFAGGCADWPEPELWEGPEGLWDERYSERHLANILGAVQMGILY			
VPEGPNGPDVAAARDIRETLARMANDEETVALIAGGHTFGTHGAGDPSFIGAE			
PEGGAIEDGLGKSGFSFGVKDAITAGLEVETWSTQPTKWSNTEFFENLAYEWELTK			
SPAGAHQWRKNAEASIPDAYEPFGKHTVMTLLDLSLRFDP IYEKISRRELPNDQF			
ADAFARWKLTHRDMGPKRVYFPELPAEDLIWQDVI PAVDHPVDKDI AELKAKV			
LATQTVQELSTWASASTFRGSDKRGANGARILAPOKDEANOPALAKVLGLV			
EGIQDFNAQTGAKKISLADILVLAGAGYKAAAGNNAVSVPLTFRMDASQAT			
DAHSFAPLRIDGFRNNGRLOFMKPEEMLVDRQAQLLTITGEMTIVVGLRLVK			
AGNPEHGVSTPEYITNDFFVNLIDVGQWVPAPRGLYRPRKTRGAAKWTGTRVD			
LIFGSHSQLRAFEVYGGADAKQKFKVDFVAANVMNADRFDLR"			
BASE COUNT	657 a	1051 c	1011 g
ORIGIN			
Query Match	22.4%	Score 333.6;	DB 1; Length 3289;
Best Local Similarity	66.8%	Pred. No. 5.8e-88;	
Matches 537; Conservative	0;	Mismatches 259;	Indels 8; Gaps 4;
QY	688	ATCCCTGGGGGCTGATTTGATTATGCC-ACCAGATTTCAACAGCTGGATGGAGGCT	746
Db	1210	ATCCGCTCGTAAGGACTTCGATATGCCGAGATATCCAAAGCTCGATCTCGACGG	1269
QY	747	CTGAAAAAGATATCAAGATTTCGTGACAACTCCAGGATTTGGTCCCTCGGATTAAT	806
Db	1270	CTGAAAAAGACTTC-ACGCGCTGATGACGGATTTCGAGGCTGGCGGCGGACTTC	1328
QY	807	GGTCATTATGGTCTTTCTTTATTCGATGGCTTGGCAGGTCGCCGAACATACAGGACA	866
Db	1329	GGTCATTATGGGCGCTCTTCATCCGATGGGCTGGCACAGCGCGGCGGACCTATCGCATC	1388
QY	867	TATGATGGCGGGGAGCGCGCTGCTGCTGACAACTGTTTGAACCTGTGAACAGCTGG	926
Db	1389	ACCGACGGCGCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	1448
QY	927	CCGGATAACGTTAATCTGGATAAGCCCTCGATTCGATTCGTGCGCCAGTCAAGAAAAATAC	986
Db	1449	CCGGACAATGCCAACCTTGACAAGGCGCGCGCTCTCTTTGGCCGATCAAGCAGAAATAC	1508

NPECPNGPDPVAAARDIRETFARMANDEETVALIAGGHTFGKTHGAGDPSFIGAD
 EGGAEIGDQKNGSTFGYGVKDAITGGPEVTWTSQTPRSNHEFFENLHNHEWELTGS
 PAGAYOGKAKNAEATIPDAYSRTSKHYPTRLTDTLSLRFDPAYEKTSRFLNEDPDEFA
 DAFARWFKLTHWASDGPYKVLGPEVPAEDLITQWDTLPVADVHLRVDETIDTAGLKVLI
 ASGLVQELVSTHMASSTFRGSDKRGANGARIRLPAQKQDWEVNRPAOLARLAKVILE
 GIORFENAOITDGGKISLADILVLAGAAVEKAAKAGGHDTITVPFTPGRMDSAEQATO
 AASFAALEPRDAGFRNVTSTROGKMGPEALVDRQAULLTLTAPETVYLGLGLRVKA
 GEPKHGVTSRPEALTNDVFNLLDMGTQWSPLEGEGBVTEGRDRTGAARWTVGTRVRD
 LIFGSHSQLRAFAEYVAQSDAREKFVKDFVAAWTKVMNADRDFLV*
 5640. 6113
 /gene= "Sma2381"
 CDS
 5640. 6113
 /gene= "Sma2381"
 /function= "Miscellaneous; Unknown"
 /note= "glimmer prediction"
 /codon_start=1
 /transl_table=11
 /product= "hypothetical protein"
 /protein_id= "AAK65945.1"
 /db_xref= "GI:14524461"
 /translation= "MLLKAMKIGRQSLDYTFGDKWLYCLAVERYSASETSAAHIALIR
 GKPRALGRIIVMERVNVNANACLGVNSICEFGQSRPDIALAHLHAADRKLKNVYVER
 IREQAQAGDLVSLUPAEVADEFLVANIAGIRIAARGGARREHLQSLSRALRAVT"
 6167. 7135
 /gene= "Sma2383"
 CDS
 6167. 7135
 /gene= "Sma2383"
 /function= "Small molecule metabolism"
 /note= "glimmer prediction; similar to TED2, a putative
 quinone oxidoreductase, from Zinnia elegans, BAA06460, and
 from cowpea, T11672; also similar to putative
 oxidoreductases from Mycobacterium tuberculosis,
 Agrobacterium tumefaciens, etc. Zinc-binding family"
 /codon_start=1
 /transl_table=11
 /product= "probable oxidoreductase"
 /protein_id= "AAK65946.1"
 /db_xref= "GI:14524462"
 /translation= "MKAVMKKEVGCTDMEFVDORPEARPAHVVEVAAAGVNFMDI
 GVROGMATIDIPNKVLGVGEGRVLAVGDTGFEFVGDVAVNYAPGSVAQROSLIPA
 ASLVKIPDVTDDRTAASTMMQGLTASHFATDFVPVPGDIALVHAAGGVGLLILTOII
 PLRGSRVIGRVSSEDKVALIARKAGAEHVIVDTDGRFADEVLRITGGEGVNVYVDGSGP
 KTFKSGLEARVSTFCWYVGLGGPGPLEIMNLPKSIKIGIYATFMDHIIHTRELLIDR
 TKQLFMIEDGSIIVTIGETVRLADAAHAAMASRATTGKLLIP*
 complement(7357. .8946)
 /gene= "Sma2385"
 CDS
 complement(7357. .8946)
 /gene= "Sma2385"
 /function= "Cell processes; Transport of small molecules"
 /note= "glimmer prediction; similar to probable ABC
 transporter from Streptomyces coelicolor, T35445, and to
 VarM (probable transporter for virginiamycin) from
 Streptomyces virginiae, BAA96297. ABC transporter family"
 /codon_start=1
 /transl_table=11
 /product= "probable ABC transporter, ATP-binding protein"
 /protein_id= "AAK65947.1"
 /db_xref= "GI:14524463"
 /translation= "MPSITLSALSWSKPDGEHVFSDLDLAFGPRTGLVGRNGIGKTS
 VLIATIGLPSTSTVIAQGRVALRILRAGADETADTVGATVAVLRAEKGDA
 SVEELTADTVSEETVISAQRALGLRADPTLLNOLSGQRTRAVLAAALHIFSPDFLL
 DEPTANVLDGRRAVIGLLSGWSSGAIIVSHDREILLEMDAIIELTSIGTKRYGGGV
 SAYOARAELEDRKASQILTKARQADVDVRKARALELRDKRQASGTRKAAKGDMPRI
 LVGRKSNAESRCKSVELARRRRAGALDAVTAAKARIEVLQFPFSIKLPTELPAQRQ
 VLAFGVTAGVDYDPAPIIRIDLSFLNKPVSVTGPNCSGKTSLLKVVTVGELGPPFPK
 VSVNPTFLDQSYSLIERGETILLENKRLNPGVSNACRAALSFRRFADALQRVE
 ALSGCVLRACLAGALGSGDPPSLIILDEPTNHLIDSIIEAVEAGLLISYDGLVWVSH
 DETFLANIGICTRVELSTSRG"
 9057. 9713
 /gene= "Sma2387"
 CDS
 9057. 9713
 /gene= "Sma2387"
 /function= "Miscellaneous; Unknown"

		/note="glimmer prediction; putative bacterial regulatory protein, tetr family"	
		/codon_start=1	
		/transl_table=11	
		/product="hypothetical protein"	
		/protein_id="AAK65948.1"	
		/db_xref="GI:14524464	
		/translation="MIRARRAGRPREEAALFRLDLSARSFTARKGIANSMEEIAEELGISKHTYRYPNPOALLEVRDLVRFKTLAEAGQCEAPLAALRDMAFVYRFGRDQYSAFYLSVAEAFVSLPRLERLAAWSAALEPLVQALISQAAGLVVPGSTIEICHVLIDLEGGANNRRLCLSPDASERLRFESRWAVFTQTMAMPEPNRPLGV"	
gene	complement(9922..10455)		
CDS	/gene="Sma2389"		
	complement(9922..10455)		
	/gene="Sma2389"		
	/function="Cell processes; Adaptation"		
	/note="glimmer prediction; very similar to Ohr, a putative stress-induced protein, from Bradyrhizobium japonicum, AAF78793"		
Query Match		22.3%;	Score 331.8; DB 1; Length 10578;
Best Local Similarity		64.9%;	Pred. No. 2.7e-87;
Matches 524; Conservative		0;	Mismatches 277; Indels 6; Gaps 2;
QY	584	TCAAATCCCTGGGGGCTGATTTTGATTTATGTCACACAGATTTCAACAGCTGGATATGGAG	743
Db	4981	TCCGATCCGCTGGGCAATACGTTCAACTACGCCGAGGAGTTCGAAGAAGCTCGACCTTCGAT	4922
QY	744	GCCTCAAAAAGCATATCAAGATTTGCTGACAACCTTCCACGAGTATGGTGCCTCGCGGAT	803
Db	4921	GCCTGAGGGGACCTCGGTGCAGCTGATGACGGATTCGAGGACTGTGTGGCGCGCGAC	4862
QY	804	TATGTCATTATGGTCTCTTTTATTTCGTATGGCTTGGCACGGTGCAGCATACAGG	863
Db	4861	TTCGGCCACTATGGCGGGCTCTTCATCCGATGGCCTGGCACAGCGCTGSCACCTACCGT	4802
QY	864	ACATATGATGGCGGGAGGCGGCAGTGTGTGCACAGCTTTTGAACGGCTGAACAGC	923
Db	4801	ATCACCGATGGCGGTGGCGCGCGCCAGGGTFCAGACGGTTCGCGCGCTCAATAGC	4742
QY	924	TGGCCGGATAACGTTAATCTGGATAAAGCCGCTCGATTCGTGTGGCCAGTCAAGAAAAA	983
Db	4741	TGGCCGGACACGCCAATCTGCACAGGCCGACGTTTGTGTGGCGGATCAACACAGAA	4682
QY	984	TACGGTCCAGTATTTCTTGGGGAGACCTGATGGTCTCTGACTGGTAAATGTGCGCTTGA	1043
Db	4681	TACGGCAACAGGATTTCTTGGCGGATCTATTGATTCACCGGCAATGTCTCGCGCTCGAA	4622
QY	1044	TCCATGGGATTTAAACGCTGGGATTTGCTGGCGGGAAGAGAACATCACTGGGAGTCGGAC	1103
Db	4621	TCCATGGGCTTCAAGACCTTCGTTTTCGGGGGCGGGTGGACGCTATGGGAACCCGAG	4562
QY	1104	CTGGTACTCTGGGGGCTTGACAAACAGCCCTTTTCAGATAACCGGGATAAAAAACG---	1160
Db	4561	GAACCTGTCTGGGACCGGAAGCACCTGCTGGCGACGAGCGCTACAGCGCGCAACGG	4502
QY	1161	AAACTTCAGAAACCTCTTGCCGCCACGACGATGGGACTATTTATGTGCAATCTCGAAGCG	1220
Db	4501	CAACTGAGCGAGCGGTTCGCGCGTGCAGATGGGTCTCATCTATTAAACCCGGAAGCG	4442
QY	1221	CCGGTGTGAAAACACAGATCTCTGGCTTCGCGGAAAGATATCAGGGAAGCTTTTTCACGT	1280
Db	4441	CCGAACGGCAACCCGACCCGGTTGTCGCGCGCGGATATCCGCAACCTTTGACCGC	4382
QY	1281	ATGGCCATGGATGATGAGGAGACTGTGGCCCTGATCGCGGAGGGGATACATTGTTGTA	1340
Db	4381	ATGCGATGAACACGACGAAGAGACCGTGGCGCTGATCGCGGTGSCACACCTTCGCGCAA	4322
QY	1341	GCACATGGTGCAGGCTCTCTGAAAATGATTTATGGCGAGGGCGCTGATGGTGCACCTGTG	1400
Db	4321	ACGCATGGCGCGCGGATCC---GTGTTTCATCTCGTGCAGACCGGAAGCGCGGCATA	4265

VGFORSVLVVTPTGSGWIDPAAMESLEVQLONGDVASVAVOYSYLASWLSLLVPEPDY
GADTAQALRDVYDIWRLPKNSRPLYLHGLSLGAMNSERSTDMDYVDFQFQALM
SGPPYKSGLRKSFTRDSFGSPAMLPFRNGSIVRFMNDQGVAAADPFVWGPRIYVL
QXASDVFDFYDLYREPDWIKPPRPDVSPELRYPPVYTFLLQTLDMAMATTAPIG
HGVAPGDIYVFGSIEVANVOGWSLQQLAELKRLHRLP
7907.8404
/gene="mlr6724"
7907.8404
/gene="mlr6724"
/note="hypothetical protein"
/codon_start=1
/transl_table=11
/protein_id="BAB52962.1"
/db_xref="GI:14026365"
/translation="MSDLVVIYVPEAKAEVRORLPAQKLEYLINISDAVIATKTEA
GNKLQMIQNTIVAGVSGFWGLLIGVIFLNPVIGVAIGAASGALGALTDVGINDA
FKKLSIATLQPNAAFLVLIHGMTADKVLISAIRKGGVGVVLRKSLDREKQIREALSQ
ASEAA"

gene

CDS

9371.10375
/gene="mlr6726"
9371.10375
/gene="mlr6726"
/note="hypothetical protein"
/codon_start=1
/transl_table=11
/protein_id="BAB52963.1"
/db_xref="GI:14026366"
/translation="WVNIATGILVLAALYFARSILMPLAFSLFVIALAWPIOMALQR
RIPRLAALITLLVIVIVIGSAIVNGELISQWLRYNTGRLOAIYASWIDWLEGH
GLAIYSPGELSFISRLVGLSHLTVRLNSVAGFAVLTFYVMLGLEAEDLQRRIRS
SLAOPGEMLSHLGRIAGKIRREVMVRSIACILGLVVAFTLVAGLELASAWAIA
FALNIPFGPLMATIPELFAIAQSGSNQALIVFVGLVNIQFPAIGSLIEPFLGAS
LAISPAVFAVFFSFMGVGPFAF IGVPILIVVVTWCAHEQSTIRWIATLLSGRST
10712.11839
/gene="mlr6727"
10712.11839
/gene="mlr6727"
/codon_start=1
/transl_table=11
/product="transcriptional regulator"
/protein_id="BAB52964.1"
/db_xref="GI:14026367"

gene

CDS

/translation="MANAPAGNHAILEPPANPAEPLVVPEOTPIPCFEFSIGD
LPPDOFAWRSGFATMLDFPQDEAGFAGTQVINDGLDALAGVSVTDSEMLLL
GVRKDPVHMLVTLIQDGRSOTITPSKSGFDAGSVQHPILGRVPEGSVTDSEMLLL
FYPRDCRMSGLAAAEFSLRGGMGPLLDYLAISLVQRLPTMGVNDLPGILSATRA
MLQVCIAQSPDHLTAHDPISATLLERARRFVQHILFQPDLSIEQMTRELIGTSRLY
RLFEASGGIVVHYIOHRLLAHAALADPNDRRLILDIADEYGFEGDGAETSFARFREFG
YCPSEVTCVKNGPSHHWHDVEVTAAPGLRLQQLRLRLQAQ"
complement(11950..12498)
/gene="mlr6729"
complement(11950..12498)
/gene="mlr6729"
/note="hypothetical protein"
/codon_start=1
/transl_table=11
/protein_id="BAB52965.1"
/db_xref="GI:14026368"

gene

CDS

/translation="MTAYENGTVSSLPPTSVRLILGLVILVGGIFVLGDLATVISA
IVIGVTAIVVGAFFEMHAFWMKGWGFIMQFLILGLIVVYVFGAMVROPVSGALVTV
IGLLVHSGVRLVFGVVRWSVGTWMLSGTFGLAGLVLITGWPSGLWVGLLLG
IDLISHGMALVIGWLPATKA"
complement(12495..15662)
/gene="mlr6730"
complement(12495..15662)
/gene="mlr6730"
/codon_start=1
/transl_table=11
/product="multidrug-efflux transport protein"

gene

CDS

Query Match 21.7%; Score 323.2; DB 1; Length 340857;
Best Local Similarity 64.6%; Pred. No. 2.4e-84;

Matches	532;	Conservative	0;	Mismatches	283;	Indels	9;	Gaps	3;
QY	672	CACAGCCCTGAATCAAAATCCCTGGGGGCTGATTTTGAATATCCACACAGATTTCACAG	731						
Db	197459	CAGTCTCTCTGTGCCACCCGATGGGTGAGCGCTTCGACTATCCCGAGGAATTCAGAGGC	197518						
QY	732	CTGGATATGGAGGCTCTGAAAAGATATCAAGATTTTCTGACAACTTCCAGGATTGG	791						
Db	197519	CTCGACCTCGACGCCCTCATCAAGGACCTGCACGCCCTGATGATCGGAGGTGG	197578						
QY	792	TGCCCTGCGGATTTATGTTCTATTTTCTTCTTCTGATGCTTGGTGGACCGTGGC	851						
Db	197579	TGCGCCGCGGATTTCCGCCACTAGCCCGCTGTTTCATCCGATGCGCTGGCACACGCCA	197638						
QY	852	GGAACATACAGGACATATGATGCCGGGGAGGCCAGTGGTGTGAGCAACCTTTTGA	911						
Db	197639	GGCACCTACCGATTGCCACGCCGCGCGCGGGCTGGCCAGCAGCGCTTTCGCG	197698						
QY	912	CGCTGAACAGCTGCCGGATACGTTAATCTGGATAAGCCCGTGCATTTGTCG	971						
Db	197699	CCGCTCAACAGCTGGCGCGCAATGTCAACCTCGACAAGCCCGCGCTTTTGTGCGCG	197758						
QY	972	GTCAAGAAAAATACGCTCCAGTATTTCCTGGGGAGACCTGATGTTCTGACTGGTAA	1031						
Db	197759	ATCAAGCAGNAATATGCCGCAAGATCTCTGGGCCGACCTTGATCTCACCGGCAC	197818						
QY	1032	GTTGCCCTTGAATCCATGGATTTAAACGCTGGGATTTGCTGCGGAGAGAGATGAC	1091						
Db	197819	GTCGCGCTGGAATCGATGGCTTCAAGACCTTTGGCTTCCGCGCGCGCGGACGTC	197878						
QY	1092	TGGGAGTCGGACCTG---GTATACTGGGGCCCTGACAACAGCCTCTTGCAGATACCGG	1148						
Db	197879	TGGGAGCCCTGAGCAGGACGCTTTACTGGGGTCCGAAAGGCAAGTGGCTGGCGCAGCGC	197938						
QY	1149	GATA---AAAAAGGAAACTTCAGAAACCTTTTCGCCCGCACGACATGGGACTATTAT	1205						
Db	197939	TACAGCGCGACCGTGCACCTGCAGAACCCGCTGCGCGCTGCAGATGGGCTGATCTAC	197998						
QY	1206	GTCAATCTCGAAGCCCGGTGAAACACAGATCTCTGCTGGCTTCCCGGAAAGATACAGG	1265						
Db	197999	GTCAATCCGGAGGACCGAAGCGCAATCCGGATCCGCTGCGCGCGCGCGATATCAGG	198058						
QY	1266	GAAGCTTTTTCAGTATGCCATGGATGATGAGAGACTGTGCCCTGATCGCGGAGGG	1325						
Db	198059	GACACATTCGCGGTATGGCGATGAACGACGAGGAAACCGTAGCACTATCGCGCGCGC	198118						
QY	1326	CATACATTTGGTAAACACATGTGCAGCTCTCTGAAATATGTTTGGCGCAGGCGCT	1385						
Db	198119	CATACGTTCCGCAAGACCCATGCT---GCGGGTGACGCGAGCCTGGTGGGTGCGAGCGG	198175						
QY	1386	GATGGTGACCTGTGGAGGACGAGGACTGGGATGAAAAATAAATGTGTACAGGAAAC	1445						
Db	198176	GAAGCGCTGATATCGAGCAGCAGGCGCTTGGCTGGCGGAGCAATTCGSCACCGCAAG	198235						
QY	1446	GGCAATATACCATCACCAGTGCCCTGGAAGGAGCCTGGTTCGAC	1489						
Db	198236	GGCGGTGACGCCCATCGGCACGCGCTGGAAGTCAATTTGGACGAC	198279						

RESULT 21

AL646061/c

LOCUS

DEFINITION

AL646061

Accession

Version

Keywords

Source

Organism

Reference

Authors

AL646061 198050 bp DNA linear BCT 07-DEC-2001
Ralstonia solanacearum GM1000 chromosome, complete sequence;
segment 5/19.
AL646061 AL646052
AL646061.1 GI:17427781
Ralstonia solanacearum.
Ralstonia solanacearum
Bacteria; Proteobacteria; beta subdivision; Ralstonia group;
Ralstonia.
1 (bases 1 to 198050)
Salanoubat,M., Genin,S., Artiguenave,F., Gouzy,J., Mangenot,S.,

Arlat,M., Billault,A., Brottier,P., Camus,J.C., Cattolico,L.,
Chandler,M., Choisine,N., Claudel-Renard,C., Cunnac,S., Demange,N.,
Gaspin,C., Lavie,M., Moisan,A., Robert,C., Saurin,W., Schlex,T.,
Sigulier,P., Thebaud,P., Whalen,M., Wincker,P., Levy,M.,
Weissenbach,J. and Boucher,C.A.

Genome sequence of the plant pathogen *Ralstonia solanacearum*

Unpublished
2 (bases 1 to 198050)

Boucher,C.A.

Direct Submission

JOURNAL

TITLE

AUTHORS

JOURNAL

Submitted (05-DEC-2001) Genoscope and CNRS UMR-8030, 2 rue Gaston
Cremieux, CP5706, 91057 Evry Cedex, France, Laboratoire de Biologie
Moléculaire des Interactions Plantes-Microorganismes INRA-CNRS,
BP27, 31326 Castanet-Tolosan Cedex, France, Fondation Jean
Dausset-CEPH, 27 rue Juliette Dodu, 75010 Paris, France, ILMG CNRS
118 Route de Narbonne, F 31062 Toulouse Cedex 4, Genoscope and INRA
URGV, 2 rue Gaston Cremieux, CP5706, 91057 Evry Cedex, France,
Laboratoire de Biometrie et Intelligence Artificielle INRA, BP27,
F31326 Castanet-Tolosan Cedex, Laboratoire de Genetique Cellulaire
INRA, BP27, F31326 Castanet-Tolosan Cedex
Christian.Boucher@toulouse.inra.fr
<http://sequence.toulouse.inra.fr/R.solanacearum.html>.

COMMENT

FEATURES

source

1. .198050
Location/Qualifiers

/organism="Ralstonia solanacearum"

/strain="GM1000"

/db_xref="taxon:305"

151. .642

/gene="RSC0770"

/note="RS05084"

151. .642

/gene="RSC0770"

/EC_number="1.-.-.-"

/function="small molecule metabolism"

/note="Product confidence : putative

Gene name confidence : hypothetical

predicted by Codon usage

predicted by Frameb"

/codon_start=1

/evidence-not_experimental

/transl_table=11

/product="PUTATIVE TRANSMEMBRANE DEHYDROGENASE (SMALL

SUBUNIT) OXIDOREDUCTASE PROTEIN"

/protein_id="CAD14472.1"

/db_xref="GI:17427782"

/translation="MTRRPSSDGDPAIGTRQMLQSLAATAASLVGSVLYAVAQ

APGEPLDAFMGLSQAALARSALDRAVGARLLAAQASAGFAAQSLAAGTLS

GLQGETALRIQAWTGVVYDGVVYEQALMGVVSILVITYCPNRPGEWAEPPVE

KQS"

642. .2261

/gene="RSC0771"

/note="RS05083"

642. .2261

/gene="RSC0771"

/EC_number="1.-.-.-"

/function="small molecule metabolism"

/note="Product confidence : putative

Gene name confidence : hypothetical

predicted by Codon usage

predicted by Homology

predicted by Frameb"

/codon_start=1

/evidence-not_experimental

/transl_table=11

/product="PUTATIVE TRANSMEMBRANE DEHYDROGENASE (LARGE

SUBUNIT) OXIDOREDUCTASE PROTEIN"

/protein_id="CAD14473.1"

/db_xref="GI:17427783"

/translation="MAETQAEQADIVVVGSGVAGALYAYELARACKSVLMLEAGPRL

PRWEIREFRQAKDMFNAPYPTFPWPHPEYGDNNYLLKGGHKFDSQYIRAVGG

TWWHAASVFLPNDLKLRSYGIARDMPIDYDLERYYGRAEALGWGPNDEBLG

SPRSQYPWAPLPSFNERTVKDALNADHPALRYVTEPVARNRPYDGRPTCCGNNNC

MPICPGAMGIVHVEKAEQAGARLIENAVVYKLEVGAKRIVAAHYKDPKGVDRHY

EKGWVFLAANGIETPKMLMSTGADFPKGVGNRSDVMGRNLMHDHPGTCVSYADRKLM
PGRQEMTSLIGRDEGFRATQAGKHLISNIRIEQETTRIFKAGKLLKPAELDAR
IRDOAARVQVDSFEHILPENRIVPSATGTDALGIPRPEITRIDDYVKRSVAVHTR
EYTAARVIGSTVQVDFHDFAPNNHITGATMGSDPKDSVVDKDCRTDFHPLFISS
SATWPTGTVNVTLTIAALRALRIADQLKEA"

2263. .3552
/gene="RSC0772"
/note="RS05082"
2263. .3552
/gene="RSC0772"
/EC_number="1.-.-.-"
/function="small molecule metabolism; energy transfer;
electron transport"
/note="Product confidence : putative
Gene name confidence : hypothetical
predicted by Codon usage
predicted by Homology
predicted by Frameb"
/codon_start=1
/evidence-not_experimental
/transl_table=11
/product="PUTATIVE OXIDOREDUCTASE DEHYDROGENASE
(CYTOCHROME C SUBUNIT) SIGNAL PEPTIDE PROTEIN"
/protein_id="CAD14474.1"
/db_xref="GI:17427784"
/translation="MRAKAASFLLLSVLGAARADADAPADAQIKRGEYLAIGD
CVACHTAPGKPFAGGLPQIPLMLGTIYSSNITDPKTKIGKWRYEERAVRGVSD
DGENLYPAMPYPSYAKINDADMDRLYAFRYGVAOVAPPPTIRWPLNRPWLKLM
NLAFLKRAPYLPKGMPTNWRGAYLAQGLAHACGCTHPRGFAMQEKAMDERCKGYLA
GSTLAGWSAYNITSDPYSGIGGWPEQIYQVLRGSGVPGQAAGPGEAVQHSFSRM
TERDVALAEYLRIVPAVSTGASRAHGWGTPATDVIALRGKPIETTDIARLYLGN
ATCHQADRGTPDGYIPLLSNLSAVGARDTNLVQVMLHGIERRAGDRHIGMPAFGRQ
LSDAALANYVTRQFGDPATPALTADEIAKRMPQ"
complement(3559. .5781)
/gene="RSC0773"
/note="RS05081"
complement(3559. .5781)
/gene="RSC0773"
/function="miscellaneous; hypothetical/global homology"
/note="Product confidence : probable
Gene name confidence : hypothetical
predicted by Codon usage
predicted by Homology
predicted by Frameb"
/codon_start=1
/evidence-not_experimental
/transl_table=11
/product="PROBABLE TRANSMEMBRANE PROTEIN"
/protein_id="CAD14475.1"
/db_xref="GI:17427785"
/translation="MRTSARVILGITPIVIVAAIVFVITFDWNRKAKPYLNDHVSQ
AVGRFAINGDLALTWKKEGSGWRAYVPMPLIASDIAIGNPDWARQPHVATVQRLL
TFVLEALPLAHRIVIPSVTIDTPMIALERDSQNRNNTWTFHQADQSQSDWKLELE
IAFSKGSVQLDEVKRIHLAATLDVDNQALYTAASGTAIRGSEPAAPASGAAASTA
PAAPDPOQYGIATVKGTYNNAAIRGSGAGGVLRLOQDARRYPLOADYVTKGTRID
LAGTLTPSSLTALDLRLHLSGSAHLYPLTVGLPDPPTDTRGLGELRQGS
WRYEQFTGRVGGSDLGTLTFAMPPECPQLTGLGVLSHQLLFADLAPLIIGADSNAS
KQRDAPVRQADKVLPEVPEFTRDWAIDADVKFTGERIVRTADLPIDHLVTHIKQ
DAVLTLDPLNFVAGGTLTNLRDQSQAPMQASAAALAAHFKIRQLFPNIESMHASI
GEINGDARLSATGNSVAALIGASNGELKILVEQGVSKFILEAMGLNVGNVILTKLFG
DKOVSINCAAGDEAVSDLAQARTFVVDODVIDVTGATSEKHESLDFTHPDSKGL
RVFSRLTPLYKGYTKHPDVSVNPAVVALRAGGAVALAPAPVNAVLPVLEQTPAPDS
PCGKLADVQRORPAPPPGKVIYRSGQTRAGDGSPPKDPAAAPARQPALPLRDPTTAGG
"

5883. .6785
/gene="RSC0774"
/note="RS05080"
5883. .6785
/gene="RSC0774"
/function="miscellaneous; hypothetical/global homology"
/note="Product confidence : probable
Gene name confidence : hypothetical
predicted by Codon usage

8622	Db	8622	GACCTCATGTCCTACGGGCAACGTCGCGCTGGAGTCGATGGGCTTCAAGACCTTTCGGC	8566
1068	QY	1068	TTTGTCTGGCGGAAGAGAAGATGACTGGGAGTCGGACCTG---GTATACTGGGGGCTGAC	1124
8562	Db	8562	TTTCGCGCGCGCGCAAGGATGTATGGGAACCGAAGAGGATGCTCTACTTGGGTTCGGA	8503
1125	QY	1125	AACAAGCCTCTTGCAGATTAACCGGGATATAAAGC---GGAACTTCAGAAACCTCTTGGC	1181
8502	Db	8502	ACCACCTGGCTGGCGCACCGAGGCTACACCGCGATCGCGATCTGGAGAAACCGCTCGCC	8443
1182	QY	1182	GCCACGCAGATGGGACTTATTATGTCAATCTCTGAAGCCCGGTCGGAACACAGATCCT	1241
8442	Db	8442	GGGTCCAGATGGGCTGTATCGTCAATCCGGAAGCCCGAAGCCCAACCCCGATCCA	8383
1242	QY	1242	CTGGCTTCGCGAAGATATCAGGGAAGCTTTTTCACGTATGGCCATGGATGATGAGGAG	1301
8382	Db	8382	ATCGCGGCGCCAGGACATCCCGAGAGCTTTCGCCCGCATGGCATGAACGATGAAGAG	8323
1302	QY	1302	ACTGTGGCCCTGATCCGGGAGGGCATACATTTGGTAAAGCACATGGTCAGGCGTCTCCT	1361
8322	Db	8322	ACGGTCGCGCTGATCGCGCGCGCCACTCTCTCGGCAAGACCCACGCGGCGCTCGGCT	8263
1362	QY	1362	GAATAATGATTGGCCAGGGCGCTGTATGTGCACTGTGGAGGAGCAGGACTGGATGG	1421
8262	Db	8262	CCAACGTG---GGCGCGGAGCGGAAGCGCGGCATCGAAGAACAGGGCTTGGCTGG	8207
1422	QY	1422	AAAAATATAATGTGGTACAGGAACGGCAATATACCATCACAGTGGCCTGGAAGAGCC	1481
8206	Db	8206	AGCAGCGCTTGGCACCGCGCAAGGCGACCGCGCATCACCAGCGGCTTGGAGTGACC	8147
1482	QY	1482	TGGTCGAC 1489	
8146	Db	8146	TGGACCAC 8139	
RESULT 22				
SRFRPS		SRFRPS	2904 bp	DNA
LOCUS		Streptomyces reticuli	furs and cpe genes.	linear
DEFINITION		Streptomyces reticuli	furs and cpe genes.	BCT 05-MAR-1999
ACCESSION		Y14317		
VERSION		Y14317.1	GI:3758888	
KEYWORDS		catalase/peroxidase; cpe gene; Fe regulatory protein; furs gene.		
SOURCE		Streptomyces reticuli		
ORGANISM		Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;		
REFERENCE		Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.		
AUTHORS		1 (bases 1 to 2904)		
TITLE		Zou,P.J., Borovok,I., de Orue Lucana,D.O., Mueller,D. and Schrempf,H.		
JOURNAL		The mycelium-associated Streptomyces reticuli catalase-peroxidase, its gene and regulation by Furs		
REFERENCE		Microbiol. 145, 549-559 (1999)		
AUTHORS		2 (bases 1 to 2904)		
TITLE		Borovok,I.		
JOURNAL		Direct Submission		
REFERENCE		Submitted (23-JUL-1997) I. Borovok, Angewandte Genetik der Mikroorganismen, FB Biologie/Chemie, University of Osnabrueck, Barbarastrasse 11, Osnabrueck D-49069, FRG		
REMARK		revised by submitter 12-Oct-1998		
COMMENT		Related entry: Y14336.		
FEATURES		Location/Qualifiers		
source		1..2904		
gene		/organism="Streptomyces reticuli"		
RBS		/strain="Tu45, described by H.Zahner, University of Tuebingen"		
		/db_xref="taxon:1926"		
		/clone_lib="pubB1"		
		/clone_lib="BamHI minilibrary"		
		78..533		
		/gene="furs"		
		78..82		
		/gene="furs"		
		/notes="putative"		

ORGANISM	Streptomyces reticuli. Streptomyces reticuli Bacteria; Firmicutes; Actinobacteria; Actinobacteridae; Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
REFERENCE	1 (bases 1 to 2904)
AUTHORS	Zou,P.J., Borovok,I., de Orue Lucana,D.O., Mueller,D. and Schrempf,H.
TITLE	The mycelium-associated Streptomyces reticuli catalase-peroxidase, its gene and regulation by Furs
JOURNAL	Microbiol. 145, 549-559 (1999)
REFERENCE	2 (bases 1 to 2904)
AUTHORS	Borovok,I.
TITLE	Direct Submission
JOURNAL	Submitted (23-JUL-1997) I. Borovok, Angewandte Genetik der Mikroorganismen, FB Biologie/Chemie, University of Osnabrueck, Barbarastrasse 11, Osnabrueck D-49069, FRG
REMARK	revised by submitter 12-Oct-1998
COMMENT	Related entry: Y14336.
FEATURES	Location/Qualifiers
source	1. .2904 /organism="Streptomyces reticuli" /strain="Tu45, described by H.Zahner, University of tuebingen" /db_xref="taxon:1926" /clone="pUBB1" /clone_lib="BamHI minilibrary" 78..533 /gene="furs" 78..82 /gene="furs" /notes="putative"
gene	
RBS	

FLNDQAFNEAFARAEKLTDRMGPKSRYIGVEPKEDLIWQDPLPOIYNPTQDI
IDLEPAIDSGLSVSELYSVASASSTRGGKRGANGARLALPQRDNDVNAAYR
ALPVLKIQKESKASLADIIVLAGVVGKAAAGLSIHVFPAPGRDARDQTDI
ENFELLEPTADGRFNRARLDVSTTESLLIDRAQQLTTPAPMTALVGMRYLGNFD
GSKNGVFTDRVGLVNDNFVLLDMRYEMKATDESKEFEGDRDRETGEVKEFTASRADL
VFGSNVLRVAEVAEYASSDAHEKFKYDFVAAMVKVMNLDREPLL*
BASE COUNT 625 a 721 c 805 g 654 t
ORIGIN About 89 minutes on K-12 map.

Query Match 20.5%; Score 305.4; DB 1; Length 2805;
Best Local Similarity 62.1%; Pred. No. 1.5e-79;
Matches 499; Conservative 0; Mismatches 301; Indels 3; Gaps 1;
QY 684 TCAATCCCTGGGGGCTGATTGTTGATGATCCACAGATTTCAACAGCTGATATGGAS 743
DB 302 TCTAACCCACTGGGTGAGGACTTTGACTACCCCAAGAATTCAGCAATAGATTACTACTAC 361
QY 744 GCTCTGAAAAAGATCAAGATTTCCTGACAACTTCCACAGATTGGTGCCTCGGGAT 803
DB 362 GGCTGAAAAAGATCTGAAGCCCTGTTGACAGATCTCAACCGTGGTGGCCAGCCGAC 421
QY 804 TATGGTCATATGTCCTTTCTTTATTCGTATGGCTGGCAGCGTCCGGAACATACAGG 863
DB 422 TGGGGCAGTTACCGCGTCTGTTTATTCGTATGGCTGGCAGCGCGGGGACTTACCGT 481
QY 864 ACATATGATGGCGGGAGCGCCAGTGTGTGTCAGCAACGTTTGAACCGCTGAACAGC 923
DB 482 TCAATGATGGAGCGGTGGCGGGTGTGTGTCAGCAACGTTTTCACCGCTGAACCTCC 541
QY 924 TGGCGGATAACGTTAATCTGGAATAAGCCGCTCGATTCTGTGGCCAGTCAAGAAAAA 983
DB 542 TGGCGGATAACGTAAGCTTCGTAATAAGCGCTGCGCTGTTGGCCATCAACAGAAA 601
QY 984 TAGGCTCCAGTATTTCTGGGGAGACCTGATGGTCTGACGTGTAATGTTGCCCTGAA 1043
DB 602 TAGGTCAGAAAATCTCTGGCGGACCTGTTTATCTCGCGGGTAACGTGGCGCTAGAA 661
QY 1044 TCCATGGGATTTAAACGCTGGATTTGCTGCGGAGAGAGATGACTGGGAGTGGAC 1103
DB 662 AACTCGGCTTCGTAACCTTCGGTTTGGTGGCGGTGTAAGACGCTGGAACCGGAT 721
QY 1104 CTGGTATACCTGGGGCTGACAAACACCTCTTGCAGATAACCGGGATAAAGCGGAAA 1163
DB 722 CTGGATGTTAAGTGGGTGATGAAGAAGCTGTGCTGACTACCGTCACTCCGAGGCTG 781
QY 1164 CTTCAAGAACCTCTTCCGCCACGAGATGGAGCTTATTTATGTCATCTCAAGGCCCC 1223
DB 782 GCGAAAGCACCGCTGGGTGCAACCGAGATGGGTCTGATTACGTTAAACCGGAGCCCG 841
QY 1224 GGTGGAACACAGATCTCTGCTTCCGGAAGATATCAGGAAGCTTTTTCACGTATG 1283
DB 842 GATCAGACGGCGAAGCCGTTCTGCGGACGAGCTATCCGCGGACCTTCGGCAACATG 901
QY 1284 GGCATGATGATGAGGAGACTGTGGCCCTGATCGGGGAGGCAATATTTGTTAAAGCA 1343
DB 902 GGCATGACAGCAGAGAACCGTGGCGCTGATGCGGGGTGATACGCTGGGTAAACCC 961
QY 1344 CATGTGTCAGCTCTCTGAAAATGTTATTTGGCGAGGCGCTGATGGTGCACCTGTGGAG 1403
DB 962 CACGTGGCG---GTCCGACATCAATGTAGTCTCTGATCCAGAAAGCTGCACCGATTGAA 1018
QY 1404 GAGCAGGACTGGGATGGAATAAATATGTTGTCAGGAAAGCGGAAATATACCATCAC 1463
DB 1019 GAACAGGTTTAGTTGGCGAGCACTTACGCGAGCGCGGCTTGGCGCAGATGCCATTACC 1078
QY 1464 AGTGGCTGGAAGGACCTGGTC 1486
DB 1079 TCTGGTCTGGAAGTAGTCTGGAC 1101

RESULT 24
AE000468
LOCUS 13840 bp DNA linear BCT 01-DEC-2000

DEFINITION Escherichia coli K12 MG1655 section 358 of 400 of the complete genome.
ACCESSION AE000468 U00096
VERSION AE000468.1 GI:1790374
KEYWORDS Escherichia coli K12.
SOURCE Escherichia coli K12.
ORGANISM Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Escherichia.
REFERENCE 1 (bases 1 to 13840)
AUTHORS Blattner, F.R., Plunkett, G. III, Bloch, C.A., Perna, N.T., Burland, V., Riley, M., Collado-Vides, J., Glasner, J.D., Rode, C.K., Mayhew, G.F., Gregor, J., Davis, N.W., Kirkpatrick, H.A., Goeden, M.A., Rose, D.J., Mau, B. and Shao, Y.
TITLE The complete genome sequence of Escherichia coli K-12
JOURNAL Science 277 (5331), 1453-1474 (1997)
PUBMED 97426617
REFERENCE 2 (bases 1 to 13840)
AUTHORS Blattner, F.R.
TITLE Direct Submission
JOURNAL Submitted (16-JAN-1997) Guy Plunkett III, Laboratory of Genetics, University of Wisconsin, 445 Henry Mall, Madison, WI 53706, USA.
Email: ecoligenetics.wisc.edu Phone: 608-262-2534 Fax: 608-263-7459
REFERENCE 3 (bases 1 to 13840)
AUTHORS Blattner, F.R.
TITLE Direct Submission
JOURNAL Submitted (02-SEP-1997) Guy Plunkett III, Laboratory of Genetics, University of Wisconsin, 445 Henry Mall, Madison, WI 53706, USA.
Email: ecoligenetics.wisc.edu Phone: 608-262-2534 Fax: 608-263-7459
REFERENCE 4 (bases 1 to 13840)
AUTHORS Plunkett, G. III.
TITLE Direct Submission
JOURNAL Submitted (13-OCT-1998) Laboratory of Genetics, University of Wisconsin, 445 Henry Mall, Madison, WI 53706, USA
COMMENT This sequence was determined by the E. coli Genome Project at the University of Wisconsin-Madison (Frederick R. Blattner, director). Supported by NIH grants HG00301 and HG01428 (from the Human Genome Project and NCHGR). The entire sequence was independently determined from E. coli K12 strain MG1655. Predicted open reading frames were determined using Genemark software, kindly supplied by Mark Borodovsky, Georgia Institute of Technology, Atlanta, GA, 30332 [e-mail: mark@ember.gatech.edu]. Open reading frames that have been correlated with genetic loci are being annotated with CG Site Nos., unique ID nos. for the genes in the E. coli Genetic Stock Center (CGSC) database at Yale University, kindly supplied by Mary Berlyn. A public version of the database is accessible (<http://cgsc.biology.yale.edu>). Annotation of the genome is an ongoing task whose goal is to make the genome sequence more useful by correlating it with other data. Comments to the authors are appreciated. Updated information will be available at the E. coli Genome Project's World Wide Web site (<http://www.genetics.wisc.edu>). *** The E. coli K12 sequence and its annotations are periodically updated; this is version M54. No sequence changes. Annotation updates: updated gene identifications and products; all new functional assignments courtesy of Monica Riley; added promoters, protein binding sites, and repeated sequences described in reference 1. The unique numeric identifiers beginning with a lowercase 'b' assigned to each gene (protein- or RNA-encoding) are now designated as gene synonyms instead of labels. This should allow them to be searched for in Entrez as gene names.

FEATURES
source
1..13840
/organism="Escherichia coli K12"
/strain="K12"
/sub_strain="MG1655"
/db_xref="taxon:83333"
complement(24..52)
/note="factor Sigma70; promoter metj3; documented +1 at 4126046"

promoter

QY 924 TGGCCGATAACGTTAATCTGATGAAGCCCGTCGATTCCTGTGGCCAGTCAAGAAAAA 983
|||||
Db 5782 TGGCCGATAACGTTAATCTGATGAAGCCCGTCGATTCCTGTGGCCAGTCAAGAAAAA 5841
|||||
QY 984 TACGGCTCCAGTATTTCTCTGGGAGACGTCGATGGTCTGACTGGTAAATGTTGCCCTTGAA 1043
|||||
Db 5842 TATGTTGACAGAAATCTCTGGGCGGACCTGTTTATCTCTCGGGTAACGTGGCGTAGAA 5901
|||||
QY 1044 TCCATGGGATTTAAACGCTGGGATTTGCTGCGGGAAGAGATGACTGGGAGTGGAC 1103
|||||
Db 5902 AACTCGGCTTCGCTACCTTCGTTTGGTGGCGGTCGTTGAAGACGTCGTGGGAACCGGAT 5961
|||||
QY 1104 CTGGTATATCTGGGCGCTGACACAGCCTCTTGACAGATAACCGGATGAAGAAACGGGAA 1163
|||||
Db 5962 CTGGATGTTAATCTGGGCTGATGAAAAGCCTTGGCTGACTCACCGTCATCCGGAAGCGCTG 6021
|||||
QY 1164 CTTCAAGAACTCTTGGCCCGCAGCAGATGGGACTTATTTATGTCATCTCTGAAGGCCCC 1223
|||||
Db 6022 GCGAAGCACCGCTGGTGCACCGAGATGGTCTGATTTAGTTAAACCGGAAGGCCCG 6081
|||||
QY 1224 GTGGGAACACAGATCTCTGCTGCTCCGCGAAGATATACAGGGAAGCTTTTTCACGTATG 1283
|||||
Db 6082 GATCACAGCGCGCAACCGCTTCTCGGCGACGAGCTATCCGCGGACCTTCGGCAACATG 6141
|||||
QY 1284 GCCATGGATGATGAGGAGACTGTGCCCTGATCGCGGAGGAGCATACATTTGGTAAAGCA 1343
|||||
Db 6142 GGCATGAACGAGGAAGAACCGTGGCGCTGATTGGCGGTGGTCATACGCTGGGTAAACC 6201
|||||
QY 1344 CATGTCGACGCTCTCTCAAAATGATTTGGCGGAGGCCCTGATGGTCACCTGTGGAG 1403
|||||
Db 6202 CACGCTGCGC---GTCCGACATCAATGTAGTCTGATCAGAGAGCTGCACCGATGAA 6258
|||||
QY 1404 GAGCAGGACTGGGATGGAATAAATGTTGATCAGGAAGACCGGAATATACCATCAC 1463
|||||
Db 6259 GAACAAGGTTAGTTGGCGAGCACTACGCGAGCGGCTTGGCGCAGATGCCATTACC 6318
|||||
QY 1464 ACTGCGCTGGAAGGACCTGTC 1486
|||||
Db 6319 TCTGGTCTGGAAGTAGTCTGGAC 6341
|||||

RESULT 25

ECOW87 96484 bp DNA linear BCT 12-AUG-2002
E. coli chromosomal region from 87.2 to 89.2 minutes.
L19201
L19201.1 GI:304961
23S ribosomal RNA; 5,10 methylenetetrahydrofolate reductase; 50S
ribosomal protein L31; 5S ribosomal RNA; 6-phosphofructokinase;
CDP-diglyceride hydrolase; DNA polymerase I; L-rhamnose isomerase;
Spot 42 RNA; aspartokinase II/homoserine; atp21II; catalase HPI;
catalase hydroperoxidase I; cdh gene; chemosensory transducer; cpxA
gene; cystathionine; cyr gene; dehydrogenase II; dsbA gene;
extragenic palindromic element; fdhB gene; fdhE gene; fdg gene;
fdhH gene; fdoI gene; formate dehydrogenase-O alpha subunit;
formate dehydrogenase-O beta subunit; formate dehydrogenase-O
gamma; gamma-synthase; glnA gene; glnG gene; glnL gene; glpF gene;
glpK gene; glpX gene; glutamine synthetase; glycerol kinase; hslU
gene; hslV gene; katG gene; kds gene; manganese superoxide
dismutase; metB gene; metF aporepressor; metF gene; metJ gene; metL
gene; mob gene; msgA gene; periplasmic sulphate binding protein;
pfkA gene; polA gene; priA gene; primosomal protein replication
factor; rhaA gene; rhaB gene; rhaD gene; rhaE gene; rhaS gene; rhaT
gene; rhamnose permease; rhamnulokinase; rhamnulose kinase;
rhamnulose-1-phosphate aldolase; rpmE gene; rfa gene; rrlA gene;
sbp gene; sodA gene; spf gene; suppressor; tpiA gene;
triosephosphate isomerase; triosephosphate mutase.
Escherichia coli.
Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
Escherichia.
1 (bases 1 to 96484)

SOURCE

ORGANISM

Plunkett, G. III, Burland, V., Daniels, D. L. and Blattner, F. R.
Analysis of the Escherichia coli genome. III. DNA sequence of the
region from 87.2 to 89.2 minutes
Nucleic Acids Res. 21 (15), 3391-3398 (1993)
93347969
8345018
This sequence was determined as part of the E. coli Genome Project
(Frederick R. Blattner, director) at the University of
Wisconsin-Madison. Supported by award HG00301 from the NIH Human
Genome Project. A preliminary report was presented at the Small
Genomes meeting, March 28-30 1993, Paris, France. The entire
sequence was independently determined from E. coli MG1655; overlaps
and conflicts with other sequence determinations are annotated. The
start of this entry overlaps the end of the entry ECOW850 (M87049)
by the three bases of an EcoRI half-site. Data kindly submitted in
computer readable form by:
Guy Plunkett III
Laboratory of Genetics
University of Wisconsin
445 Henry Mall
Madison WI 53706
USA
Phone: 608-262-2534
ecoli@genetics.wisc.edu
Email: 608-263-7459.
Fax: Location/Qualifiers
1. .96484
/organism="Escherichia coli"
/strain="K-12"
/sub_strain="MG1655"
/db_xref="taxon:562"
/cissue_lib="lambda"
/note="This sequence comprises the following lambda
clones: EC21-16, EC27-1225, EC27-649, EC19-119, EC23A58,
EC27-267, EC18-126. M13mp19 or Janus vectors were used for
subcloning."
1. .19281
/note="corresponds to lambda clone EC21-16"
1. .2063
/gene="rrlA"
1. .2063
/gene="rrlA"
/product="23S ribosomal RNA"
/note="CG Site No. 203; 23S rRNA"
1. .3
/gene="rrlA"
/note="Eco RI half-site is 3bp overlap with end of M87049
ECOW850"
2157. .2276
/gene="trfA"
2157. .2276
/gene="trfA"
/product="5S ribosomal RNA"
/note="CG Site No. 210; 5S rRNA"
2282. .2321
/note="putative"
2459. .2483
/note="putative"
2504. .2540
/standard_name="REP; repetitive extragenic palindromic
element."
/note="contains 1 REP sequence"
complement(2546..3058)
/standard_name="ORF_f170"
/note="matches PS00017: ATP_GTP_A"
/codon_start=1
/transl_table=11
/protein_id="AA02991.1"
/db_xref="GI:304962"
/translation="MIPLAFAAWSGTGKTLKKLPALCARGIRPGIKKTHHMD
VDKPKDSYELKAGAAOTIVASQORWALMTETPDEEELDQFLASMDYTSKLDLILY
EGFKHEETAKIVLFRDGAHRPEELVIDRHVAVASDVPLNDVALLINDVEGLADF
VVEWMQKQNG"

CDS

repeat_region

terminator

terminator

rRNA

gene

misc_feature

rRNA

gene

misc_feature

rRNA

gene

misc_feature

rRNA

gene

misc_feature

rRNA

gene

misc_feature

rRNA

gene

misc_feature

rRNA

gene

misc_feature

rRNA

gene

misc_feature

rRNA

gene

misc_feature

rRNA

gene

misc_feature

rRNA

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 REMARK
 FEATURES
 source

Philipps-Universitaet-Marburg, Marburg, Germany
 2 (bases 1 to 3018)
 Scherer,M.
 Direct Submission
 Submitted (25-JAN-2001) Scherer M., Biochemistry,
 Max-Planck-Institut fuer Terrestrische Mikrobiologie,
 Karl-von-Frisch-Str. 35043 Marburg, GERMANY
 revised by author 10-MAY-2001

location/Qualifiers
 1. 3018
 /organism="Aspergillus nidulans"
 /db_xref="taxon:5072"
 393. 2612
 /gene="cpeA"
 393. 2612
 /gene="cpeA"
 /EC_number="1.11.1.6"
 /codon_start=1
 /product="catalase-peroxidase"
 /protein_id="CAC59821.1"
 /db_xref="GI:15384983"
 /translation="MGSNECPYSRONANTGGGNNRDMPPDLKLNLRHNSVNP
 LDRGFFDTAFNSLDYFGLKRDLEALMTDSQDWWPADFGHYGLFIRMAHWSAGTVRV
 FDRGGGGOQORFAPLNSWPNVDSQFVGGKEWLNDRVLYNGELNPLAASHMLIYV
 ESMGFTEGAPAGSRTWEADQSVFVGGKEWLNDRVLYNGELNPLAASHMLIYV
 NPSPKNNPDLVLAARDIRITFGMAANDEETVALLAGHTFGKTHGAGPATHLGKEP
 HGAGIELQGLWGESGESGTHAITSGLIEVITKPTKWSNOFFLEYLYDWELTKS
 PAGAHQYVAKGVPEFFIPDPFDPSPKHPRLMTDLSLYDPEYERISRRLENPDQFA
 DAFARAVKLTFRDGVPRVLYQGPEVSEVLIWQDPVPLDHPVIDNDIATLTKKAIL
 NSGISHTDLSTAWASASTFRGSKRGANGARIRLSPKWNKVNQSPWLSLALE
 KIOQNDAQSDTKRSLADLIYLAGAASLEAKARDAGHNVSFPPGRDADQEOFD
 VDSFNLEIADGFNRYGRCTPVLTEDELIDKAQLNLNLSPELTVLIGLRLVNNY
 DRNIGVTFARQLNDFVNLNLDGMQVMPADDTNEIFIGSDRTGTQARWAKSRAD
 LVFGSHAEIRLRAISEVSSDGEAKFYKDFVAWEKVSNLDRLFKTGLAQRIKPOL"
 BASE COUNT 660 a 883 c 789 g 686 t

Query Match 20.4%; Score 303.4; DB 8; Length 3018;
 Best Local Similarity 61.2%; Pred. No. 6.2e-79;
 Matches 528; Conservative 0; Mismatches 326; Indels 9; Gaps 2;

QY 632 CTACTATCCAGAACACTGGATTAACTCTCTGAGATTACACAGCCCTGAATCAATCC 691
 DB 464 CTGTTGCCGCGAGACTCAACTGACATCCCTCTGTCAGCACAACTCGTCTTAATCC 523
 QY 692 CTGGGGGCTGATTTGATTATGCGACAGATTTCAACACTGGATGAGAGGCTCTGAA 751
 DB 524 CCTGGATAAGGCTTTGACTATCTATGCGCGGTTCAATAGCTCGACTACTTCGAGCTGAA 583
 QY 752 AAAAGATATCAAGATTTGCTGACAACTCCAGGATTTGTCCTCGGATTTATGGTCA 811
 DB 584 GCGGGATCTGGAGGCACTATGACAGACTCCAGGACTGGTGGCGGCGGACTTTGGTCA 643
 QY 812 TTATGTCCTTTCTTTATCGTATGGTTGGCAGGTCGCGGAACATACAGGACATATGA 871
 DB 644 CTATGGGGAATCTTTATCGCATGGCTGGCAGACTGCTGGACAGTATCGGCTTTGA 703
 QY 872 TGGCCGGGAGGCGGCGGAGTGGTGTGACAGACGCTTTTGAACCGCTGAACAGCTGGCGGA 931
 DB 704 CGGTGCGGCGGTGGCGGAGGCTGACGACGCTTCGCTCCGCTCAACAGCTGGCGGA 763
 QY 932 TAAGCTTAATCGATTAAGCCCGTCTGATTTGCTGTGCGCAGTCAAGAAAAATACGCTC 991
 DB 764 TAAGCTGAGCTTGGCAAGGCTCGTCTCTTTATGGCCCATCAAGCAGAGTACGGAAG 823
 QY 992 CAGTATTCTTGGGAGACCTGATGGTCTGACTGGTGAATGTTGCCCTTCAATCCATGGG 1051
 DB 824 CAGATCTCAATGGCTGATTTGATCTTCGCGGGAATGTCGCCCTTGAATCAATGGG 883
 QY 1052 ATTTAAACGCTGGGATTTGCTGGCGGAGAGAGATGACTGGGAGTCGGACCTGGTATA 1111
 DB 884 GTTCAAGACCTTTGGGTTTGGCGGTGCGCGAGTATGATACCTGGGAGCAGACCACTGGT 943

QY 1112 CTGGGGCCTGACAAACAA-----GCCTCTTGAGATAACCGGATATAAAACGGGAAACT 1165
 DB 944 CTTCTGGGAGCGGAGAGGAATGGTTGGTAAATGATGTCCGTACTTGAACGGAGAACT 1003
 QY 1166 TCAGAAACCTCTTTCGCCGCCACGAGATGGGACTATTATATGTAATCTCAATCCTCAAGCCCGG 1225
 DB 1004 CGACACCCGCTTTCGCCGCATCACATGGTCTTATTTAGTTTATCCAGAGACCCAA 1063
 QY 1226 TGGAAACACAGATCCTCTTCCCGGAAAGATATCAGGAAAGCTTTTTCAGCTATGGC 1285
 DB 1064 CAAGAACCCGACCCGCTTCTCGCGGCAAGGATATCCGCATCACCTTTGTGCAATGGC 1123
 QY 1286 CATGGATGATGAGAGACTGTGGCCTGATCGGGGAGGAGCATACATTGGTGAAGCACA 1345
 DB 1124 CATGAATGATGAGAGACTGTGGCCTGATTTGGTGGACACACGTTTCGGAAGACGCA 1183
 QY 1346 TGGTGCAGCTCTCTGAAAAATGATTGGCGGAGGCTTGTGGTGCACCTGTGGAGGA 1405
 DB 1184 CGG---CGCGGCGCTTCAACCCATCTCGCAAGAAGACACATGGTGGGTTATGAGTT 1240
 QY 1406 GCAGGACTGGGATGAAAAATAATATGTGTACAGGAAACGCAATATACCATCACCAG 1465
 DB 1241 ACAAGGCTTAGCTGGGAGAGCGCTTTCGAGTCTGGGACCGGCGGACATCTATCACCAG 1300
 QY 1466 TGGCCTGGAAGGAGCCTGGTCTGA 1488
 DB 1301 CGCTCTGAGGTGATCTGGACCA 1323

RESULT 28
 LOCUS AE005625
 DEFINITION Escherichia coli O157:H7 EDL933 genome, contig 3 of 3, section 244 of 230.
 ACCESSION AE005625 AE005174
 VERSION AE005625.1 GI:12518851
 KEYWORDS
 SOURCE Escherichia coli O157:H7 EDL933.
 ORGANISM Escherichia coli O157:H7 EDL933
 Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Escherichia.

REFERENCE 1 (bases 1 to 10667)
 AUTHORS Perna,N.T., Plunkett,G. III, Burland,V., Mau,B., Glasner,J.D., Rose,D.J., Mayhew,G.F., Evans,P.S., Gregor,J., Kirkpatrick,H.A., Posfai,G., Hackett,J., Klink,S., Boutin,A., Shao,Y., Miller,L., Grobeck,E.J., Davis,N.W., Lim,A., Dimalanta,E., Potamouisis,K., Apodaca,J., Anantharaman,T.S., Lin,J., Yen,G., Schwartz,D.C., Welch,R.A. and Blattner,F.R.
 TITLE Genome sequence of enterohaemorrhagic Escherichia coli O157:H7
 JOURNAL Nature 409 (6819), 529-533 (2001)
 MEDLINE 21074935
 PUBMED 11206551
 REFERENCE 2 (bases 1 to 10667)
 AUTHORS Perna,N.T., Plunkett,G. III, Burland,V., Mau,B., Glasner,J.D., Rose,D.J., Mayhew,G.F., Evans,P.S., Gregor,J., Kirkpatrick,H.A., Posfai,G., Hackett,J., Klink,S., Boutin,A., Shao,Y., Miller,L., Grobeck,E.J., Davis,N.W., Lim,A., Dimalanta,E., Potamouisis,K., Apodaca,J., Anantharaman,T.S., Lin,J., Yen,G., Schwartz,D.C., Welch,R.A. and Blattner,F.R.
 TITLE Direct Submission
 JOURNAL Submitted (22-OCT-2000) Laboratory of Genetics, University of Wisconsin, 445 Henry Mall, Madison, WI 53706, USA
 FEATURES
 source
 1. 10667
 /organism="Escherichia coli O157:H7 EDL933"
 /strain="EDL933"
 /serotype="O157:H7"
 /db_xref="taxon:155864"
 /note="enterohemorrhagic"
 <1. .251
 /note="O-island #162: Region of the EDL933 chromosome not homologous to E. coli K-12 MG1655"
 complement(253. .861)

misc_feature
 gene

```

/gene="yix"
/notes="25492"
complement(253..861)
/gene="yix"
/function="orf; Unknown function"
/notes="Residues 1 to 202 of 202 are 97.52 pct identical to
residues 1 to 202 of 202 from Escherichia coli K-12 Strain
MG1655: B3937"
/codon_start=1
/transl_table=11
/product="orf, hypothetical protein"
/protein_id="AAG59138.1"
/db_xref="GI:12518852"
/transl_table="MKRLILSLVSPAFAMQPQTGDIIFQISRSSQSKAIQLATH
SDYSGTMLVNRNKKPVYFAVGKYPVPLKQWIAHGEKGVYVRRVGGSLVQEQOK
LAQAKRYLGPYDFSPSWSDRQCYCEVVKVYQNALGMVGEQQLKEFDLSNPLV
QAKLKERYGNIPLEEVVSPQAFDAPQLTVAKWPLFSW"
complement(921..1238)
/gene="metJ"
/notes="25493"
complement(921..1238)
/gene="metJ"
/function="regulator; Amino acid biosynthesis: Methionine"
/notes="Residues 1 to 105 of 105 are 100.00 pct identical
to residues 1 to 105 of 105 from Escherichia coli K-12
Strain MG1655: B3938"
/codon_start=1
/transl_table=11
/product="repressor of all met genes but metF"
/protein_id="AAG59139.1"
/db_xref="GI:12518853"
/transl_table="MAEWSGEYISPYAEHKKSEVKKIVTSIPLKVLKILTDERRR
QVNLRIATNSLCEALFHAFTGQPLDDADLRKERSDEIPEAAKEIMREMGINPET
WEY"
1515..2675
/gene="metB"
/notes="25494"
1515..2675
/gene="metB"
/function="enzyme; Amino acid biosynthesis: Methionine"
/notes="Residues 1 to 386 of 386 are 99.48 pct identical to
residues 1 to 386 of 386 from Escherichia coli K-12 Strain
MG1655: B3939"
/codon_start=1
/transl_table=11
/product="cystathionine gamma-synthase"
/protein_id="AAG59140.1"
/db_xref="GI:12518854"
/transl_table="MTRKQATIAVRSGLNDEOYGVVPPHLSSTYNTFGENEPRAH
DYSRGNPTRVQRALEGGAGVLTNTGMSAHLVTYFLKPGDLLVAHPCYV
GSLRFLDSLRKRGYRVLPDQDEQALRAALNPKLVLPESPNLLRVDIADIC
HLAREVAVVVDNTFLSPALQNLGADLVLSCTKYLNGHSDVVGVIADPDV
VTELAWANNITVVGGAFFDSYLLRLGLTLPVRLMELAQNAKAIYVLOQPLVKLY
HPSLPENQGHIAARQOKGFGAMLSFELDGDQTLRFLGLSLFLAESLIGVESLI
SHAATMHAGNAPARAAGISEILLRISTIGIEDGDLIADLNGFPRAANK"
2678..5110
/gene="metI"
/notes="25495"
2678..5110
/gene="metI"
/function="enzyme; Amino acid biosynthesis: Methionine"
/notes="Residues 1 to 810 of 810 are 99.87 pct identical to
residues 1 to 810 of 810 from Escherichia coli K-12 Strain
MG1655: B3940"
/codon_start=1
/transl_table=11
/product="aspartokinase II and homoserine dehydrogenase
II"
/protein_id="AAG59141.1"
/db_xref="GI:12518855"
/transl_table="MSVIAQAGAKRQLHFKFGSSLDVRCYKLRVAGIMAEYSQDDM
MVVASAGSTTNLNLKLSQTLDSAHQVQTLRRYQCDLISGLLPAEADSLISAF
VSDLERLAALDSDINDAVYAEVVGHGVNSRLMSAVLNQOGLPAAWLDAREFLRAE

```

```

RAAQPVQDEGLSYPLQLLQVHPGKRLVVTGFIISRNAGETVLLGRNGSDYSATQIG
ALAGYSRTVITWSDVAGVTSADPRKVDACLPLLRIDEASELARLAAPVHLARTQLPV
SGSEIDLQLRCSYTPDOGSTRIERYLAGSTGARIVTSHDDVCLIFQVPTSDQFKLAH
KEIDQLKRAQVRPLAVGVNDROLLOPCYTSFVADSALKILDEAGLPGELRLQGLA
LVAMGAVYTRNPLCHRWQOLKQOPVEFTWQSDGSLVAVALTGTPTSLQSLHQ
SVFRAKRLGLVLFGRGNIGSRWLELFAREQSTLSARTGFEFVLAVGVVDSRRSLISYD
GIDASRALAFNDNEVEODESLFLMMRAHPYDVLVDVLTASOQLADQYLDFAHSGF
HYTSANKLAGASDSNKYRIQIHDAFKTRHLYNATVGAGLPINHTVRDLIDSGDTIL
SISGIPSGTSLWLFQDGSVPFTLVDAQOQGLTEPDRDLGKDVMRKLVILAR
EAGYNTEPDQVRVESLVPACFEGSIDHFEFENGDELNEQVORLEAAAREMGVLVRVA
REFDANGKARVGVEAVREDHPLASLLPCDNVFALESRYWYRDNPLVIRGPGAGRDVTAGA
IOSDNLRLAQLL"
5459..6349
/gene="metF"
/notes="25496"
5459..6349
/gene="metF"
/function="enzyme; Central intermediary metabolism: Pool,
multipurpose conversions of intermed. met.m"
/notes="Residues 1 to 296 of 296 are 99.66 pct identical to
residues 1 to 296 of 296 from Escherichia coli K-12 Strain
MG1655: B3941"
/codon_start=1
/transl_table=11
/product="5,10-methylenetetrahydrofolate reductase"
/protein_id="AAG59142.1"
/db_xref="GI:12518856"
/transl_table="MNFHASORDALNQSALAEVQOQINVSFEFFPPTSEMBOITLWNS
IDRLSLKPKFVSVTVYGANSRDRTHSIKIKORTGLEAAPHLCIDATPDELRTI
ARDYNNGIRHIVALRGDLPGSGKPEYASDLVLLKEVADFEDLSVARYPEVHPEAK
SAQADLLNKRKVADAGANRAITOFFFVSVYLRDRDRCVSAGIDVEIIPGLPVSFNRK
QAKKFADMTNVRIPAMWAQMFQDGLDDDAETRLKLVGANIAMDVMVKILSRGVKDFHYT
LNRAEWSYAICHTLGVVRPL"
6678..8858
/gene="katG"
/notes="25497"
6678..8858
/gene="katG"
/function="enzyme; Protection responses: Detoxification"
/notes="Residues 1 to 726 of 726 are 99.86 pct identical to
residues 1 to 726 of 726 from Escherichia coli K-12 Strain
MG1655: B3942"
/codon_start=1
/transl_table=11
/product="catalase; hydroperoxidase HPI(I)"
/protein_id="AAG59143.1"
/db_xref="GI:12518857"
/transl_table="MSTSDDIHNTATGKCPFHGGHGDOSAGAGTTTRDWPNNOLRVD
LLNQHNSRNPGLGEDDYRKESKLDYGLKDKLKALTESOPWMDWGSYAGLFIIR
MAWHGATYRSIDGRGAGRGQRFAPLNSWPDNVSIDKARLLIPKIKYQKXISWA
DPLFAGNVALENSGRTFGGAGREDVWEPDLVNWGDEKAWLTHRHPEALAKAPLG
ATEMGLIYVNEPDPHSGSPLSAAAIATFNGMNDDEETVALLTAGHTLKGTHGAG
PTSNVAPDEAAP IEEQGLGWASTYSGVGADAITSGLEVWVQTPTQWSNYFFENLF
KEMWOTRSPAGAIQEAVDPEIIPDPDPKSRKPTMLVDTLRFDPPEKISRR
FLNDPQAFNEAFARAFKLTDRMDGPKSRVIGPEVPKEDLIWQDPLPQIYNTPTODI
IDLKFAIDSGLSVSELVSVANASSTFRGGKRGANGARLALMPQRDWDVNAARVR
ALPVELEKIQKESKASLADILVLAGVVGVEKASAAAGLSIHVPFAPRVDARQDDTDI
EMFLELPTADGFRNTYRLADLSTTESLIDKAAQTLTAPETALVGMVIRGANFDI
GSKNGVTDRTVGVLSDNFNLDIMRYENKATDESEKELFEGDRDRETEGVKYSASRADL
VFGSNSVLRAVAVYASSDAHEKFKVDFVAWVKVNNLDRFDLL"
8919..9857
/gene="yijE"
/notes="25498"
8919..9857
/gene="yijE"
/function="orf; Unknown function"
/notes="Residues 1 to 312 of 312 are 98.71 pct identical to
residues 1 to 312 of 312 from Escherichia coli K-12 Strain
MG1655: B3943"
/codon_start=1
/transl_table=11
/product="orf, hypothetical protein"
/protein_id="AAG59144.1"

```

[illegible]

Db 1779 CGCTTCCGCCCTTGAACAGCTGGCGGACACACCAATCTCGACAGGCCGCCGCTG 1838

QY 963 CTGTGGCCAGTCAAGAAAAAATAGCGCTCAGATATTCCTGGGAGACCTGATGTCCTG 1022

Db 1839 CTGTGGCCAGTCAAGCAGAAAGTACGGCGCCAAAGCTGTCTGGCGCCAGCTCTATGTCGTG 1898

QY 1023 ACTGGTAATGTTGCCCTTGAATCATCATGGGATTTAAACACGCTGGGATTTGCTGGCGGAAGA 1082

Db 1899 GTCGGCACAGCTGGCCCTTGGAGTCGATGGCTTCAAGACCTTCGGCTTCGGCGCGCGCCG 1958

QY 1083 GAAGATGACCTGGGAGTCGGACCTGGTATCTACTGGGGCCCTGACACAGCCCTCTTGGCAGAT 1142

Db 1959 GCCGACCACTGGGAGCCCGAAGAGCTGTACTGGGGTCCCGAAAGCAGCTGCTTGGACGAC 2018

QY 1143 AACCGGATAAAGACG---GGAACCTTCAGAAACCTCTTCGGCCGACGAGATGGGACTT 1199

Db 2019 AAGCGCTACAGCGCGGCGCTGAGCTGGACTCGCCCTGGCGCGGTCCAGATGGGGCTG 2078

QY 1200 ATTATGTCAATCTCTGAAGCGCCCGGTGGAACACAGATCCTCTGGCTTCGGCGAAGAT 1259

Db 2079 ATCTACGTCATCTCCGAGGCGCCGACGCAACCCGACCCGCTGGCTTCGGCGCGGAC 2138

QY 1260 ATCAGGGAAGCTTTTACGATATGGCCATGATGATGAGAGACTGTGGCCCTGATGCGG 1319

Db 2139 ATTCGAGACCTTCGCGCGGATGGCCATGAATGACGAGGAGACCGTCGCGCTGATGCC 2198

QY 1320 GGAGGACATATTTGGTAAAGACATGTTGTCACGCTCTCTGAAAAATGTTATGGCGCA 1379

Db 2199 GCGGCCACACTTCGCGCAAGCTCTACGG---CGCGGCGACGCTCCTCTCGTGGCGTT 2255

QY 1380 GGGCTGATGTGTCACCTGTGGAGGAGCAGGACTGGGATGGGAAAAATAATGTTGATCA 1439

Db 2256 GAGCGGGAAGCGCGCGGATCGAGGCCCGAGGCTTGGTGGCGGACGACGACGCGACG 2315

QY 1440 GGAACCGCAATATACCATCAGTGGCTTGAAGGAGCTGCTGTC 1486

Db 2316 GCAAGGCGCGCGGCGGATCAGCGGCGGCGGCGGAGTGATCTGGAC 2362

RESULT 30

AE005967

LOCUS

DEFINITION

AE005967 Caulobacter crescentus CB15 section 293 of 359 of the complete genome.

ACCESSION

AE005967

VERSION

AE005967.1

KEYWORDS

GI:13424686

SOURCE

Caulobacter crescentus CB15

ORGANISM

Caulobacter crescentus CB15

REFERENCE

1 (bases 1 to 10677)

1 Nierman, W.C., Feldblyum, T.V., Laub, M.T., Paulsen, I.T., Nelson, K.E., Eisen, J., Heidelberg, J.F., Alley, M.R., Ohta, N., Maddock, J.R., Potocka, I., Nelson, W.C., Newton, A., Stephens, C., Phadke, N.D., Ely, B., Deboy, R.T., Dodson, R.J., Durkin, A.S., Gwinn, M.L., Haft, D.H., Kolonay, J.F., Smit, J., Craven, M.B., Khouri, H., Shetty, J., Berry, K., Utterback, T., Tran, K., Wolf, A., Vamathevan, J., Ermolaeva, M., White, O., Salzberg, S.L., Venter, J.C., Shapiro, L. and Fraser, C.M.

Complete genome sequence of *Caulobacter crescentus*

Proc. Natl. Acad. Sci. U.S.A. 98 (7), 4136-4141 (2001)

21173698

11259647

2 (bases 1 to 10677)

1 Nierman, W.C., Feldblyum, T.V., Paulsen, I.T., Nelson, K.E., Eisen, J., Heidelberg, J.F., Alley, M.R., Ohta, N., Maddock, J.R., Potocka, I., Nelson, W.C., Newton, A., Stephens, C., Phadke, N.D., Ely, B., Laub, M.T., Deboy, R.T., Dodson, R.J., Durkin, A.S., Gwinn, M.L., Haft, D.H., Kolonay, J.F., Smit, J., Craven, M., Khouri, H., Shetty, J., Berry, K., Utterback, T., Tran, K., Wolf, A., Vamathevan, J., Ermolaeva, M., White, O., Salzberg, S.L., Shapiro, L., Venter, J.C. and Fraser, C.M.

TITLE

JOURNAL

FEATURES

source

1. .10677

/organism="Caulobacter crescentus CB15"

/strain="CB15"

/db_xref="taxon:190650"

220. .2433

/gene="CC3043"

220. .2433

/gene="CC3043"

/note="identified by match to TIGR protein family HMM TIGR00198"

/codon_start=1

/transl_table=11

/product="catalase/peroxidase"

/protein_id="AAK25005.1"

/db_xref="GI:13424687"

/translation="METPARRNTPMKAVEDNIAGCPMGHGRGPNRDMWPQSLRLE GLNQHPRNPMEAFDYAEAFKSLDLDAVVDLHALMTDSQEWPAADFHYGGGLFIR LAWHAAGTVRTDGRGAGGQORFAPLNSWPDNTNLDKARLLWPIKQYKAGLSWA DYLVLNVALESMGPKTFAGGADOWPEPELYWGPSTWLDKDRKSGERELDSPL GAVQMLIYNDEGNGNDPLASARDIRETFARMANDEETVALIAGTFTGKAHA GDASLVGPEGAIEAQFGWASKHGCTGKGDALITGGEVITWTOTPTFWSNHFENL FRYEWELTSPAGAKOWAKNADIPDAPFNKTHVPRMLTSDLRALRDPAYEKISR REYENDQFADAFARAFWKLTTRDMGPPIGRYLGPIVPRKEELIWDQPIPAVDHPLADDK DAALKAKATLGLSASDLVSAWASATYRQSKRGKGANGARIRLAKQDWANVNP VLAKVLALEGVQKDFNASAGGKTSILADLIVLAGAAIEAKAADAGTSVTPFPAG RMDASAEOTDAHSFEALPRSDGRNRYGPKHNAPEALVDRAQLLGLSGELTVL VGGRLVLANADGSKDVTNRPGALSNDFFNLLSMETTWSPSTANAFAGHDKRSE PRWATRVDLIFGSHAEALRPAFEVYACASQEKFCVDFVTANKVYNADRLDLAA"

complement(2685..4538)

/gene="CC3044"

complement(2685..4538)

/gene="CC3044"

/note="identified by match to TIGR protein family HMM TIGR00110"

/codon_start=1

/transl_table=11

/product="dihydroxy-acid dehydratase"

/protein_id="AAK25006.1"

/db_xref="GI:13424688"

/translation="MPPYRSRTTTHGRNMAGRLWRATCMKEDFGKPIIAVANSTF QVPGHVHLKDLQGLVAREAGGVAKENTIAVDGIAHGGMGLYSLPSRLDIAD SVETMVAHCADAIVCISNCDKITPCMLNARMNLNIPVYVSGGPMACKVTVKGIIR ALDLDANVVAADDSYDEVEAIEAKAECTCGSCSMFTANSNCLTEALGLSLPGN GSVLATHADREALFKEAGRVVVDLCORWYEQDATALPRGIATRAAFENASLDIANG GSTNTVHLIAAAHEGIDFSMADIDRLSHVPCLSKVAPEAKSDHVEDVHRAGVMA ILGELERGLIDASQPTVHAPTMEALARDWIDRTNSQIAHEFFKAAPGGKPTQVAFS QAAWBELEDRENGVIRSVHPFSDKGLAVLGNLAPGECIVKTAGVDESILTPRG TARVFSODAAVSGLLGGOVKAGEVVVIRYEGPKGGQMEGLYPTTYLKSGLGAC ALVTDGRFSGTSGLSGHSVSPAGGGLIALVETGDPILIDITPTGILTLESDAVLA ARREAQLARGKDANTPLNKRDLTPALRAAAMTTNAARGAVRDVDSQIERG"

4637. .5527

/gene="CC3045"

4637. .5527

/gene="CC3045"

/note="identified by match to protein family HMM"

/codon_start=1

/transl_table=11

/product="isoleucine biosynthesis transcriptional activator"

/protein_id="AAK25007.1"

/db_xref="GI:13424689"

/translation="MDIRQFNPFAVAETHFGGAERLGIOTPLSQSQLEAKALG APFATKRHVHVELTALQKWLPHVLEALAAVDALPDPARRLRDQGTLSLSFVSTAD YSLPDLVRYTAEAFPGVEIQLVDEATSDVQVPAIQAGERHAGIIPPNKSLPAALAY RRLVSEPLVAVTPEAAGEPLDLAALADLPVLFPRTVAPAFHDLVTGVAAQGPV RIVQEATQMTIILISVSAGLGMALAPASLRKARVGRVYVDLVDPPIILERGLVWRDE AAPTLQGLILRTEDGSAPD"

5578. .6093

/gene="CC3046"

QY	688	ATCCCTGGGGGGCTGATTTTGATATATCCACAGATTTCAACAGCTGGATATGAGGCTC	747
Db	167	ACCCGATGGGTGGGGCTTGCAGTATGCGCGGAGTTCGCGACCATCGAGCTTGACGCC	226
QY	748	TGAAAAAGATATCAAGATTTCTGACAACTTCCAGGATTCGTCCTCGCGGATATG	807
Db	227	TGACGGGACATCGAGGAGTATGACACCTCGAGCGCTGTGGCCGCGCAGTACG	286
QY	808	GTCAATATGCTCTTTCTTTATTCGTATGCGGCTTGGCAGCGTCCGGAACATACAGACAT	867
Db	287	GCCACTACGGCGCTGTTTATCCGATGCGGTGGCAGCTGCGGCACTACCGCATCC	346
QY	868	ATCATGGCGGGAGGCGGACGCTGCTGGTCAGCAAGCTTTGAACCGCTGAACAGCTGC	927
Db	347	ACGACGGCGCGCGCGCGGGCGGATGACGCGGTTCGCGCGCTTAACAGCTGC	406
QY	928	CGGATAACGTTAATCTGGATAAAGCCGCTGATTCGTGGCCAGTTCGTGGCCAGTCAAGAAAAATACG	987
Db	407	CCGACAAACGCGCTTGGACAGCGCGCGCTGCTGTGGCCGCTCAAGAAAGTACG	466
QY	988	GCTCCAGTATTCCTGGGAGACCTGATGCTCTGACTGTGTAATGTTGCCCTTGAATCCA	1047
Db	467	GCAAGAAGCTCTCATGGCGGAGCTGATTTCTTTTCGCGGCACTGCGCGCTGGAATCGA	526
QY	1048	TGGGATTAACAGCTGGGATTTGCTGGCGGAGAGAGATGACTGGGAGTCCGACCTGG	1107
Db	527	TGGGCTCAAGACGCTCGGCTTCGGCTTCGCGCGGTCGACCATGGGAGCGCGATGAGG	586
QY	1108	TATACTGGGGGCGCTGACAAAGCCCTTTCAGATAACCGGGATA--AAAACGGGAAC	1164
Db	587	TCATTTGGGCAAGGAGCCACCTGGCTGGCGATGAGCGTTACAGCGGTAAACGGGATC	646
QY	1165	TTCAGAACTCTTGGCGGACAGCAGATGGGACTTATTTATGTCATCTCTGAAGGCCCG	1224
Db	647	TGAGAAACCGCTGGCGCGGTGCAGATGGGCTGATCTACGTGAACCGGAGGCGCA	706
QY	1225	GTGAAACAGATCTCTGCTTCCGCGGAAATATCAGGGAAGCTTTTTCAGTATGG	1284
Db	707	ACGGCAACCGGAGCCCATGGCGCGCGGTGACATTCGCGAGAGCTTTTCGCGCATGG	766
QY	1285	CCATGATGATGAGGAGCTGTGCGCTGATCCGCGGAGGCGCATACATTTGTAAGACAC	1344
Db	767	CCATGAGACGCTCGAAACAGCGCGCTGATCTCGCGGTACACTTTTCGTTAAGACC	826
QY	1345	ATGGTCAGGCTCTCTGAAAATGATTTGGCGAGGCGCTGATGTCACCTGTGAGG	1404
Db	827	ATGGCGCG--GCCCGCGGATCTGCTCGCGCCGCAACCCGAGGCTGCTCCGCTGAGC	883
QY	1405	AGCAGGACTGGGATGGAAAAATATGTTGGTACAGGAACGCAATATACCATCACCA	1464
Db	884	AGATGGCTTGGGCTGGAAGAGCTGCTATGGCACCGGACCGGTAAAGGAGCGATCACCA	943
QY	1465	GTGGCTGGAAGGAGCTGTGTCCA 1488	
Db	944	CGGCAATCAGGCTCGTATGACCA 967	
RESULT 33			
AF314115			
LOCUS			
DEFINITION			
Mycobacterium tuberculosis isolate F16 catalase-peroxidase (katG)			
gene, complete cds.			
AF314115			
VERSION			
AF314115.1 GI:12744169			
KEYWORDS			
SOURCE			
ORGANISM			
Mycobacterium tuberculosis.			
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;			
Actinomycetales; Corynebacterineae; Mycobacteriaceae;			
Mycobacterium; Mycobacterium tuberculosis complex.			
1 (bases 1 to 223)			
REFERENCE			
AUTHORS			
Orru, G., Iona, E., Memmi, G., Oggioni, M.R., Fattorini, L., Orefici, G.			
and Pozzi, G.			
Mutation Associated to Isoniazid Resistance in Italian Isolates of			
Mycobacterium tuberculosis			
Unpublished			
REFERENCE			
2 (bases 1 to 223)			
AUTHORS			
Orru, G.			
Direct Submission			
Submitted (18-OCT-2000) Dip. Biologia Molecolare, Laboratorio di			
Microbiologia Molecolare e Biotecnologia, Università di Siena,			
Viale Bracci, Siena 53100, Italy			
JOURNAL			
FEATURES			
source			
1..223			
/organism="Mycobacterium tuberculosis"			
/isolate="F16"			
/db_xref="taxon:1773"			
/note="isoniazid resistant isolate"			
gene			
1..223			
/gene="katG"			
1..223			
/gene="katG"			
/codon_start=1			
/note="KatG"			
/transl_table=11			
/product="catalase-peroxidase"			
/protein_id="AAK06517.1"			
/db_xref="GI:12744170"			
/translation="MPEQHPITETTTTGAASNGCPVVGHNKYPVEGGNODWPNRLN			
LKVHONPAVDPMGAADYAAEVATIDVALTRDIEVNTTSOPWMPADYGHVPLF			
IRMAHAAATYRIHGRGAGGCMORFAPLNSWPDNASLDKARLLWPVKYKGGKLS			
WADLIYPACNALESMEKFTFGFCGFRVDQWEDVYWGKEATWLDERSGRDLN			
PLAIVOMGLIYVNPENPGNPDMAAARDIREFRMAMNDVETAAALVGGHTFGKTH			
ILYGYELTKSPAGWQYAKDAGAGTTPDFGFGSRPTMLATDLSLRVPIYER			
ITRRWHEPELAEFAKAWYKLIHRDMGVARYLGPLVKQTLWDPVPVPOYSHDLV			
GEAEIASLSOIRASGLTVSOLVSTAWAASFRGSKRGANGRLQPOYGVN			
DPDGLKVRITLLEIOESNSAAGNIKVSFADLVLSGCAAEKAAKAAGNITVP			
FTPTRDASQBDQDVFVLEPKADCFRNLGKNPLPAEYMLDKKLLATLSAPEM			
TVLVGGLRVIGANTKRLIPGFTFEASESLTNDFVNLDMGITWEPSPADDTGYGKD			
GSGKVKWTSRVLFGVSGNSELSRLVEYGADDAQPKFVQDFVAANDKVNMLDRFDVR			
BASE COUNT			
427 a 695 c 742 g 359 t			
ORIGIN			
Query Match			
20.08; Score 298.4; DB 1; Length 2223;			
Best Local Similarity 62.4%; Pred. No. 1.8e-77;			
Matches 502; Conservative 0; Mismatches 296; Indels 6; Gaps 2;			
QY	688	ATCCCTGGGGGGCTGATTTTGATATGCGCGGAGTTCGCGACCATCGAGCTTGACGCTC	747
Db	167	ACCCGATGGGTGGGGCTTGCAGTATGCGCGGAGTTCGCGACCATCGAGCTTGACGCC	226
QY	748	TGAAAAAGATATCAAGATTTCTGACAACTTCCAGGATTCGTGGCCCTCGCGGATATG	807
Db	227	TGACGGGACATCGAGGAGTATGACACCTCGAGCGCTGTGGCCGCGCAGTACG	286
QY	808	GTCAATATGCTCTTTCTTTATTCGTATGCGGCTTGGCAGCGTCCGGAACATACAGACAT	867
Db	287	GCCACTACGGCGCTGTTTATCCGATGCGGTGGCAGCTGCGGCACTACCGCATCC	346
QY	868	ATCATGGCGGGAGGCGGACGCTGCTGGTCAGCAAGCTTTGAACCGCTGAACAGCTGC	927
Db	347	ACGACGGCGCGCGCGCGGGCGGATGACGCGGTTCGCGCGCTTAACAGCTGC	406
QY	928	CGGATAACGTTAATCTGGATAAAGCCGCTGATTCGTGGCCAGTTCGTGGCCAGTCAAGAAAAATACG	987
Db	407	CCGACAAACGCGCTTGGACAGCGCGCGCTGCTGTGGCCGCTCAAGAAAGTACG	466
QY	988	GCTCCAGTATTCCTGGGAGACCTGATGCTCTGACTGTGTAATGTTGCCCTTGAATCCA	1047
Db	467	GCAAGAAGCTCTCATGGCGGAGCTGATTTCTTTTCGCGGCACTGCGCGCTGGAATCGA	526
QY	1048	TGGGATTAACAGCTGGGATTTGCTGGCGGAGAGAGATGACTGGGAGTCCGACCTGG	1107

Db 527 TGGGCTTCAAGACGTTCCGGTTCGGCTTCGGCCGGTCCGACCACTGGAGCCCGCATGAGG 586
 QY 1108 TATACCTGGGGGCTTGACAAACAGCCTCTTGACAGATAACCGGGATA---AAAACGGGAAAC 1164
 Db 587 TCTATTGGGGCAAGAACCCACCTGGCTCGGCGATGAGCGTTACAGCGGTAAGCGGATC 646
 QY 1165 TTCAGAAACCTCTGCGCCACGACGATGGGACTTATTATGTCATCCCTGAAGGCCCGC 1224
 Db 647 TGGAGAACCCGCTGGCGCGGTGCAGATGGGGCTGATCTAGTGAACCCGAGGGGCCGA 706
 QY 1225 GTGAAAACACAGATCTCTGCTTCCGCGAAAGATATCAGGAAGCTTTTTCACGTATGG 1284
 Db 707 ACGCAACCCGACCCCATGCCCCGGGCTGCACATTCGGAGACGTTTCGGGCGATGG 766
 QY 1285 CCATGGATGATGAGAGACTGTGCCCTGATCGCGGAGGCGATACATTTGGTAAAGCAC 1344
 Db 767 CCATGAACGAGCTGAAACAGCGCGCTGATCGTCCGCGCTCACACTTTCGGTAAGACC 826
 QY 1345 ATGTGACGCTCTCTGAAATGATTTGGCGCAGGCGCTGATGTGACCTGTGGAGG 1404
 Db 827 ATGGCGCG---GCCCGCGGATCTGTGCGCCCGCCGAACCCGAGGCTGCTCCGCTGGAGC 883
 QY 1405 AGCAGGACTGGATGGAATAAATCTGTACAGGAAACGGCAATATACCATCACCA 1464
 Db 884 AGATGGGCTTGGCTGGAAGAGCTGATGGCACCGAACCGGTAGGACGCGATCACCA 943
 QY 1465 GTGGCTGGAAGAGCCTGGTCCA 1488
 Db 944 CGGCATCGAGGTCGTATGGACGA 967

RESULT 34

AF314117

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

SOURCE

gene

CDS

AF314117 2223 bp DNA linear BCT 11-FEB-2001
 Mycobacterium tuberculosis isolate F15 truncated
 catalase-peroxidase (katG) gene, complete cds.

AF314117.1 GI:12744173

Mycobacterium tuberculosis.
 Mycobacterium tuberculosis.

Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 Actinomycetales; Corynebacterineae; Mycobacteriaceae;

Mycobacterium tuberculosis complex.
 1 (bases 1 to 2223)

Orru, G., Iona, E., Memmi, G., Oggioni, M. R., Fattorini, L., Orefici, G.
 and Pozzi, G.

Mutation Associated to Isoniazid Resistance in Italian Isolates of
 Mycobacterium tuberculosis

Unpublished
 2 (bases 1 to 2223)

Orru, G.
 Direct Submission

Submitted (18-OCT-2000) Dip. Biologia Molecolare, Laboratorio di
 Microbiologia Molecolare e Biotecnologia, Università di Siena,
 Viale Bracci, Siena 53100, Italy

Location/Qualifiers
 1..2223

/organism="Mycobacterium tuberculosis"
 /isolate="F15"

/db_xref="taxon:1773"

/note="isoniazid resistant isolate"

1..2223
 /gene="katG"

1..84
 /gene="katG"

/note="katG"

/codon_start=1

/transl_table=11

/product="truncated catalase-peroxidase"

/protein_id="AAK06519.1"

/db_xref="GI:12744174"

/translation="MPEQHPPIETTTTGAASNGCPVVGHMK"

BASE COUNT 429 a 694 c 741 g 359 t

ORIGIN

Query Match 20.0%; Score 298.4; DB 1; Length 2223;

Best Local Similarity 62.4%; Pred. No. 1.8e-77;

Matches 502; Conservative 0; Mismatches 296; Indels 6; Gaps 2;

QY 688 ATCCCTGGGGGGTGATTTTGTATGATGCCACACAGATTTCAACAGCTGGATATGAGGCTC 747

Db 167 ACCGATGGTGGGGTTCGACTATGCCGCGAGGTCCGACCATCGACTTACGCTTACGCC 226

QY 748 TGAATAAGATATCAAGATTTGCTGACAACTTCCAGGATTTGTCCTTCGGATATG 807

Db 227 TGACGCGGACATCGAGGAAGTATGACCACTCGCAGCCGTGTGGCCCGCGACTAGC 286

QY 808 GTCAATATGCTCTTCTTTTATTCGTATGCTTGGCACGGTCCCGAATACACAGCAT 867

Db 287 GCACCTACGGCGCGCTGTTTATTCGGATGGCTGGCAGCTCCCGCACTACCGCATCC 346

QY 868 ATGATGGCGGGAGGCGCAGTGTGTCAGCAACAGTTTGAACCGCTGAACAGCTGCG 927

Db 347 ACGACGGCGCGCGCGCGCGCGCGCATCGAGGTTCCGCGCGCTTAACACCTGCG 406

QY 928 CGGATAACGTTTATCTGGATAAAGCCGCTCGATGCTGTGGCCAGTCAAGAAAAATACG 987

Db 407 CCACAAACCCAGCTTGGACAAGGGCGCGCGCTGCTGTGGCGGCTCAAGAAGATAGC 466

QY 988 GCTCCAGTATTTCTGGGAGACCTGATGCTCTGCTGCTGCTGCTGCTGCTGCTGCTG 1047

Db 467 GCAAGAGCTCTCANTGGCGGACCTGATGTTTTCGCGGCACTGCGCGCTTGAATCGA 526

QY 1048 TGGGATTTAAACGCTGGGATTTGCTGGCGGAAGAGAAGTACTGGAGTGGACCTGG 1107

Db 527 TGGGCTTCAAGACGTTTCGGGTTTCGGCTTCGGCGCGGTCGACCGAGTGGGAGCCG 586

QY 1108 TATCTGGGCGCTGACACAAAGCCTCTTCGAGATAACCGGGATA---AAACGGGAAAC 1164

Db 587 TCTATTGGGGCAAGAACCCCTGCTGGCGGATGAGCGTTTACAGCGGTAAAGCGGATC 646

QY 1165 TFCAGAAACCTCTTCCCGCCAGCAGATGGGACTTATTATGTCATCTCTGAAGCCCG 1224

Db 647 TGGAGAACCCGCTGGCGCGGTGACAGTGGGCTGATCTAGTGAACCCGAGGGGCGCA 706

QY 1225 GTGAAAACAGATCCTCTGGCTTCCGCGAAGATATCAGGGAAGCTTTTTCACGTATGG 1284

Db 707 ACGGCAACCCGACCCCATGCGCGCGGCTGACATTCGCGAGACGTTTCGGCGCATGG 766

QY 1285 CCATGATGATGAGGAGACTGTGGCCCTGATCGCGGAGGCGCATACATTTGGTAAAGCAC 1344

Db 767 CCATGAACGCTCGAAGACGCGCGCTGATCTGCGCGGTACACTTTCGGTAAGACCC 826

QY 1345 ATGGTGCAGCGTCTCTGAAAAATGATTTGGCGAGGCGCTGATGCTGACCTGTGGAGG 1404

Db 827 ATGGCGCG---GCCCGCGGATCTGTCGCGCCCGGACCCGAGGCTGCTCCGCTGAGC 883

QY 1405 ACCAGGAGCTGGATGGAATAAATAATGTTGTAAGAAACGCAATATACCATCACCA 1464

Db 884 AGATGGGCTTGGGCTGGAAGAGCTGCTATGGCACCGGAACCGGTAAAGACGCGATCACCA 943

QY 1465 GTGGCTGGAAGAGCCTGGTCCA 1488

Db 944 GCGCATCGAGTCTGATGGACGA 967

RESULT 35

AF314120

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

AF314120

2223 bp

DNA

linear

BCT 11-FEB-2001

Mycobacterium tuberculosis isolate An05

catalase-peroxidase (katG)

gene, complete cds.

AF314120

AF314120.1

GI:12744179

Mycobacterium tuberculosis.

ORGANISM	Mycobacterium tuberculosis
REFERENCE	Bacteria; Firmicutes; Actinobacteria; Actinobacteridae; Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium tuberculosis complex.
AUTHORS	1 (bases 1 to 2223) Orri, G., Iona, E., Memmi, G., Oggioni, M.R., Fattorini, L., Orefici, G. and Pozzi, G.
TITLE	Mutation Associated to Isoniazid Resistance in Italian Isolates of Mycobacterium tuberculosis
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 2223) Orri, G.
AUTHORS	Direct Submission
TITLE	Submitted (18-OCT-2000) Dip. Biologia Molecolare, Laboratorio di Microbiologia Molecolare e Biotecnologia, Università di Siena, Viale Bracci, Siena 53100, Italy
JOURNAL	Location/Qualifiers
FEATURES	1..2223 /organism="Mycobacterium tuberculosis" /isolate="Au05" /db_xref="taxon:1773" /note="isoniazid resistant isolate" 1..2223 /gene="katG" 1..2223 /gene="katG" /note="katG" /codon_start=1 /transl_table=11 /product="catalase-peroxidase" /protein_id="AAK06522.1" /db_xref="GI:12744180"
source	/translation="MPEQHPPTETTTTGAASGCPVVGHHKYPVEGGNODWPNRLN LKVLHONVADPMPGAADYAAEVATIDVADLTDEEYMTT SQPMPADYGHYPLF IRMAHAAGTYRIHDCGGAGGQMRFAPLNSPDNASLDKARLLPWPVKYKKLS WADLIVFAGNCALESMEFTFGFGRVQWDEPDEVYKKEATLWDERYSKGRDLEN PLAAVOMGLIYNPEGPNPNPMAAVDIREFREMANNDVETALILVGHTEGKTH GAGPADLVPEPEAPLEQKGLWKSXYGTGTGKDAITSGTIEVWNTWTFKNDNFLE ILGYEWEUTKSPAGAWQYAKDGAGATIPDPFGPGRSPTMLADLSLRVDPIYER ITRRWLEPELADERAKAWYKLIHRDMPGVARYLGPLVKQTLWDDVPVAPSHDLV GEATIASLKSQIRASGLTWSQLVSTAAWASPRGSDKRGANGRRIRLPQVGVWVN DPGDRLGILRTLEEIOESFNSAAGNIKVFIADVLVGGCAIEKAAKAAAGHNITVP FTPGRDASQEDTDESVFAVLEPKADGFNVLGKGNPLPAEYMLDKANLLTSAPEM TVLVGLRVLGANYKKRLPLGVFTFEASESLTNFFNLDGMTWESPADDDGTGYQKD GSGKVKWTSRVDLVVGSNSLRVVEVYGADDAQPKFVQDFVAAWKNVNDLRFQVR "
BASE COUNT	428 a 695 c 742 g 358 t
ORIGIN	
Query Match	20.0%; Score 298.4; DB 1; Length 2223;
Best Local Similarity	62.4%; Pred. No. 1.9e-77;
Matches	502; Conservative 0; Mismatches 296; Indels 6; Gaps 2;
QY	688 ATCCCTGGGGCGCTGATTTGATATGATCCACAGATTTCACAGCTGGATGAGGCTC 747
Db	167 ACCCGATGGTGGCGGCTTCGACTATGCCGCGGAGTCCGACCATCGAGCTTGACGCC 226
QY	748 TGAATAAAGATATCAAGATTGCTGACAACTTCCAGGATTGTCGCCCTCGCGATTATG 807
Db	227 TGACGGGACATCGAGGAGTATGACACCTCGACGCGTGTGGCCGCGGACTACG 286
QY	808 GTCAATTATGTCCTTTCTTTTATTCGTATGCTTGGCAGCTGCCGGAACATCAGGACAT 867
Db	287 GCCACTACGGCCGCTGTTTATCCGGATGCGTGGCAGCTGCCGCGACCTACCGCATCC 346
QY	868 ATGATGCCGGGAGGCGGACGATGGTGTGATCAGCAAGCTTTTGACCGCTGAACAGTGGC 927
Db	347 ACGACGCCCGCGCGCGCGGCGGCGGATCGAGCGGTTCGCGCGCTTAACAGCTGGC 406
QY	928 CGGATACGCTTAATCTCGATAAAGCCGCTGATGCTGTGCGCAGTCAAGAAAAATACG 987
Db	407 CCGACACGCCGCTTGACAGCGCGCGGCTGCTGTGCGCGGTCAAGAAAGTACG 466
QY	988 GCTCCAGTATTTCTTGGGGAGACCTGATGCTCTGACTGGTATGTTGCCCTTGATCCA 1047
Db	467 GGAAGAAGCTCTCATGGCGGACCTGATGTTTTCGCGGCACTCGCGCTGGAATCGA 526
QY	1048 TGGGATTTAAACCGCTGGGATTTGCTGGCGGAAGAAGATGACTGGAGTCGAGCTGG 1107
Db	527 TGGGCTTCAAGACGTTTCGGGTTTCGGCTTCGCGCGGTTCGACCACTGGAGCCGATGAGG 586
QY	1108 TATATCGGGGCTTGACAAACAGCCCTTTGACAGATAACCGGGATA---AAAAGGGAAAC 1164
Db	587 TCTATTGGGGCAAGGAAGCCACCTGGCTCGCGATGATGAGGCTTACAGCGGTAAAGCGGATC 646
QY	1165 TTCAGAAACCTTTGCGCCGACCCAGATGGGACTTTATTTATGTCATCTTGAAGGCCCGC 1224
Db	647 TGGAGAACCGCTTGGCGCGGTGCAGATGGGCTGATCTACGTGAACCGGAGGCGCGCA 706
QY	1225 GTGGAACACAGATCTCTGGCTTCGCGGAAGATATCAGGGAAGCTTTTTCACGTATGG 1284
Db	707 ACGGCAACCGGACCCCATGCGCGCGGTTCGACATTTCGAGAGAGCTTTTCGCGCATGG 766
QY	1285 CCATGATGATGAGGAGACTGTGGCCCTGATCGCGGAGGCGCATATTTGGTAAAGCAC 1344
Db	767 CCATGAACACGCTCGAAACAGCGCGCTGATCTCGCGGCTCACACTTTTCGGTAAAGACC 826
QY	1345 ATGGTCAGGCTCTCTGAAAATGTTATGGCCAGGCGCTGATGTCGACCTGTGTGAGG 1404
Db	827 ATGGCGCG---GCCGCGCGATCTGTGCGGCGCGCAACCGAGGCTGTCCCTGTGAGC 883
QY	1405 AGCAGGAGCTGGGATGGAATAATAATGTGTACAGGAACGCAATATACCATCACCA 1464
Db	884 AGATGGCTTGGCTGGAAGAGCTCGTATGGCACCGGACCGGTAAAGCGCATCACCA 943
QY	1465 GTGGCCTGGAAGGAGCTGTGTCGA 1488
Db	944 GCGGCATCGAGTCTGTATGACGA 967
RESULT	36
AF314121	2223 bp DNA linear BCT 11-FEB-2001
LOCUS	Mycobacterium tuberculosis isolate Rm09 catalase-peroxidase (katG)
DEFINITION	gene, complete cds.
ACCESSION	AF314121
VERSION	AF314121.1 GI:12744181
KEYWORDS	
SOURCE	Mycobacterium tuberculosis.
ORGANISM	Mycobacterium tuberculosis Bacteria; Firmicutes; Actinobacteria; Actinobacteridae; Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium tuberculosis complex. Orri, G., Iona, E., Memmi, G., Oggioni, M.R., Fattorini, L., Orefici, G. and Pozzi, G.
REFERENCE	1 (bases 1 to 2223)
AUTHORS	Mutation Associated to Isoniazid Resistance in Italian Isolates of Mycobacterium tuberculosis
TITLE	Unpublished
JOURNAL	2 (bases 1 to 2223)
AUTHORS	Orri, G.
TITLE	Direct Submission
JOURNAL	Submitted (18-OCT-2000) Dip. Biologia Molecolare, Laboratorio di Microbiologia Molecolare e Biotecnologia, Università di Siena, Viale Bracci, Siena 53100, Italy
FEATURES	Location/Qualifiers 1..2223 /organism="Mycobacterium tuberculosis" /isolate="Rm09" /db_xref="taxon:1773" /note="isoniazid resistant isolate" 1..2223 /gene="katG" /gene="katG" 1..2223 /gene="katG" /note="katG"

[illegible]

```

Db 167 ACCGATGGTGGCGGCTTCGACTATGCCGGGAGGTCGCGACCATCGACGTTGACGCC 226
Qy 748 TGAATAAGATATCAAGATTGCTGCAACTTCCCAAGGATGGTGCCTCGCGGATATG 807
Db 227 TGACGGGAGCATCGAGAGTAGTATACCACTCGACGCGTGTGCGCGCGGACATGAC 286
Qy 808 GTCATATGTCCTTCTTATTCGATGCTTGGACGCTGGCGGACATACAGGACAT 867
Db 287 GCCACTACGGCGCGCTTTATCCGATGGCGTGGACGCTGCGCGGACCTACCGCATCC 346
Qy 868 ATGATGCGCGGGAGGCGCAGTGGTGCAGCAACGTTTGAACCGCTGAACAGCTGGC 927
Db 347 ACGACGCGCGCGCGCGCGGCGCATGCAGCGTTCGCGCGCTTAAACAGCTGGC 406
Qy 928 CGGATACGTTATCTGGATAAAGCCGTCGATGCTGTGGCGGACGTAAGAAAATACG 987
Db 407 CCGACAAACGCGAGTTGGACAAGCGCGCGGCTGTGTGGCGGCTCAAGAAGATACG 466
Qy 988 GCTCCAGTATTTCTGGGAGACCTGATGCTGCTGACTGTAATGTTGCCCTTGAATCCA 1047
Db 467 GCAAGAAGCTCTCATGGCGGACCTGATGTTTTCGCGCGCACTGCGCGCTGGAATCGA 526
Qy 1048 TGGGATTTAAACGCTGGGATTTGCTGGCGGAAGAGATGACTGGGAGTTCGGACCTGG 1107
Db 527 TGGGCTTCAAGAGCTTCGGGCTTCGGCTTCGCGCGGTCGACAGTGGGAGCCCGATGAG 586
Qy 1108 TATACGGGCGCTGCACAACAGCCTTTCGAGATAACCGGGATA----AAACGGGAAC 1164
Db 587 TCTATTGGGCAAGGAAGCACCTGGCTCGCGCATGAGCGTTACAGCGGTAAAGCGGATC 646
Qy 1165 TTCAAGAACCTCTTGGCGCCACGACGATGAGCTTATTTATGCAATCTTGAAGGCCCG 1224
Db 647 TGAGAACCCGCTGGCGCGCGGTCAGATGGGCTGATCTACGTGAACCCGAGGCGCGGA 706
Qy 1225 GTGGAACACGATCTCTCGCTTCGCGGAAGATATCAGGGAAGCTTTTTCACGATGG 1284
Db 707 ACGCAACCCGAGCCCATGCGCGCGCGTGCACATTCGCGAGACGTTTCGCGCATGG 766
Qy 1285 CCATGGATGATGAGGAGCTGTGGCCCTGATCGCGGGAGGCGATACATTTGGTAAGAC 1344
Db 767 CCATGAACAGCTCGAAGACGCGCGCTGATCGTGGCGGTGCACACTTCGCGTAAGACC 826
Qy 1345 ATGCTGACGCTCTCTGAAATGTTATTCGCGCGCGCTGATGTCGACCTGTGAGG 1404
Db 827 ATGCGCGG--GCCTGGCGATCTGTGTCGCGCCGCGCAACCCGAGGCTGCTCGCGTGG 883
Qy 1405 AGCAGGACCTGGGATGGAATAAATGCTGTACAGGAACGCGCAATATACCATCACCA 1464
Db 884 AGATGGCTTGGGCTGGAAGAGCTCGTATGGCACCGGAACCGGTAAGGACGCGATCACA 943
Qy 1465 GTGGCTTGGAGGAGCCTGTGCGA 1488
Db 944 CGGCAATCGAGGTCGTATGACGA 967

```

RESULT 38

MTU40595

LOCUS

DEFINITION Mycobacterium tuberculosis H251/90, linear BCT 07-JAN-1996
catalase-peroxidase (katG) gene, complete cds.

ACCESSION U40595

VERSION U40595.1

KEYWORDS GI:1150699

SOURCE Mycobacterium tuberculosis strain=H0251/90.

ORGANISM

Mycobacterium tuberculosis
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
Actinomycetales; Corynebacterineae; Mycobacteriaceae;
Mycobacterium; Mycobacterium tuberculosis complex.

1 (bases 1 to 2223)

Marttila.H.J., Soini.H., Huovinen.P. and Viljanen.M.K.

katG gene mutations in isoniazid-resistant Mycobacterium

tuberculosis-strains isolated from Finnish patients

JOURNAL Unpublished

REFERENCE

AUTHORS Marttila.H.J., Soini.H., Huovinen.P. and Viljanen.M.K.
TITLE Direct Submission

JOURNAL

Submitted (13-NOV-1995) Harri J. Marttila, Mycobacterial Reference
Laboratory, National Public Health Institute, Kiinamyllynkatu 13,
Turku, SF-20520, Finland

FEATURES

source

1..2223

/organism="Mycobacterium tuberculosis"

/strain="H0251/90"

/db_xref="taxon:1773"

/note="INH-resistant strain"

1..2223

/genes="katG"

1..2223

/genes="katG"

/codon_start=1

/transl_table=11

/product="catalase-peroxidase"

/protein_id="AA085169.1"

/db_xref="GI:1150700"

/translations="MPBOHPITETTTTGAASNGCPVVGHMKYPVGGGNDWPNRLN
LKVLHONPAVDPMGAADYAAEVAFTDIDALTPDIEVNTSQPWADYGHGPLE
IRAWHAGTYRIHDGCGAGGGMORFAPLNSPNDNASLDKARLLMPVKKYGKLLS
WADLVFAGLCALSMGFKTFGFGVQDWEPEDEVYMGKEATWIGDERYSGRKDLN
PLAAVQMGLLIYNPEEGPNPNPMAAADIRETFRMANNDVETAAALIVGHTFGKTH
GAGPADLVGPEPAALQOMGLGWKSGYGTGDKAITSIEVWNTPTKWDNSFLE
ILXYEWELTKPAGAWOYAKDGAGATIPDPGGPSRPTMLATDLSLRVDPYIER
ITRWLHPELADFEAKYKLIHRDMPVARYLGLVPKQLLWQDPVPVASHDLV
GEAELSLKSOIRASGLTVSOLVSTAAWASFGSGDKRGANGGRIRIQPGWVEYN
DPDGLURKVIITLEEIPESNSAAGNIKVSFADLVVLGCAAEKAAKAAAGHNTVP
FTFGRDASQEDTVESFALEPKADGRNLYLGRNPLDPAEYMLLDKANLLTSAPEM
TVLVGGLRVLGANYKRLPLGVFTFEASESLNDFVNLMDGITWEPSPADGTQGVKD
GSKVKWTKSRVDLVFGSNSELRALVEVYGAADDAQPKFVQDFVAAMKVNMLDREDVR
"

BASE COUNT 427 a 696 c 741 g 359 t

ORIGIN

Query Match 20.0%; Score 298.4; DB 1; Length 2223;
Best Local Similarity 62.4%; Pred. No. 1.8e-77;
Matches 502; Conservative 0; Mismatches 296; Indels 6; Gaps 2;

Qy 688 ATCCCTGGGGGCTGATTTTGTATGATCCACGAGTTTCAACAGCTGGATGAGGCTC 747
Db 167 ACCGATGGTGGCGCTTCGACTATCCGCGGAGTTCGCGACCATCGACGTTGACGCC 226

Qy 748 TGAATAAGATATCAAGATTGCTGACAACTTCCGAGGATTCGCGCTCGCGATATG 807
Db 227 TGACGCGGACATCGAGGAAGTATGATACCCTCGCAGCGTGGTGGCGCGCATACG 286

Qy 808 GTCATATGTCCTTTCTTTTATGCTATGGCTTGGCAGCTTCGCGGAACATACAGGACAT 867
Db 287 GCCACTACGGCGCGCTGTTATCCGATGGCGTGGCAGCTTCGCGGACCTACCGCATCC 346

Qy 868 ATGATGCGCGGGAGGCGCGCAGTGGTGCAGCAACCTTTTGAACCGCTGAACAGCTGGC 927
Db 347 ACGACGCGCGCGCGCGCGGCGGCATGACGCGGTTTCGCGCGCTTAAACAGCTGGC 406

Qy 928 CGGATACGTTAATCTCGATAAAGCCGTCGATGCTGTGGCCAGTTCGAAGAAAATACG 987
Db 407 CCGACAACGCCAGCTTGGACAAGCGCGCGGCTGCTGTGGCGGCTCAAGAAGATGACG 466

Qy 988 GCTCCAGTATTTCTGGGAGACCTGATGGTTCCTGACTGGTAAATGTTGCCCTTGAATCCA 1047
Db 467 GCAAGAAGCTCTCATGGCGGACCTGATGTTTTCGCGGCACTGCGCGCTGGAATCGA 526

Qy 1048 TGGGATTTAAACGCTGGGATTTGCTGGCGGAAGAGATGACTGGGAGTTCGGACCTGG 1107
Db 527 TGGGCTTCAAGAGCTTCGGGTTTCGGCTTCGCGCGGTCGACAGTGGGAGCCCGATGAG 586

Qy 1108 TATACGGGCGCTGCACAACAGCCTCTTCGAGATAACCGGGATA---AAACGGGAAC 1164

```

Db 587 TCTATTGGGCAAGAAAGCCACCTGGCTCGCGCATGAGCGTTACAGCGTTAAACGGGATC 646
QY 1165 TTCAAAACCTCTTCCGCCGACAGATGGGACTTATTTATGTCAATCTCTGAAGGCCCG 1224
Db 647 TGGAGAACCCGCTGGCGCGGTGCAGATGGGCTGATCTACGTGAACCCGGAGGCCGA 706
QY 1225 GTGAAACACAGATCTCTGCTTCCGCGAAAGATATCAGGAAGCTTTTTCACGTATGG 1284
Db 707 ACGCAACCCGACCCATCGCGCGCGGTGCACATTCGCGAGAGCTTTCGCGCATGG 766
QY 1285 CCATGGATGAGGACACTGTGCCCTGATCGCGGAGGCGATACATTTGGTAAAGCAC 1344
Db 767 CCATGAAGAGCTGGAACAGCGCGCTGATCTGCGCGCTACACTTTCGGTAAGACCC 826
QY 1345 ATGTGCGAGGCTCTCTCTGAAATATGATTCGCCAGCGGCTGATGTGCACTGTGGAGG 1404
Db 827 ATGGCGCG--GCCCGCGGATCTGTGTCGCCCGCGAACCAGGCTGCTCCGCTGGAGC 883
QY 1405 AGCAGGACTGGGATGGAATAATATGTGTGTCAGGAACCGCAATATACCATCACCA 1464
Db 884 AGATGGCTTGGGCTGGAAGAGCTCGTATGGCACCGGAACCGTAAGGAGCGCATACCA 943
QY 1465 GTGGCTTGAAGAGCGCTGTGCGA 1488
Db 944 CGGCATCGAGTCTGTATGACGA 967

RESULT 39
LOCUS MTU41305 2223 bp DNA linear BCT 07-JAN-1996
DEFINITION Mycobacterium tuberculosis INH-resistant strain H0359/91, mutant
ACCESSION U41305
KEYWORDS catalase-peroxidase (katG) gene, complete cds.
SOURCE U41305.1 GI:1150703
ORGANISM
    Mycobacterium tuberculosis strain-H0359/91.
    Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
    Actinomycetales; Corynebacterineae; Mycobacteriaceae;
    Mycobacterium; Mycobacterium tuberculosis complex.
REFERENCE
    1 (bases 1 to 2223)
    Marttila,H.J., Soini,H., Huovinen,P. and Viljanen,M.K.
    katG gene mutations in Isoniazid-resistant Mycobacterium
    tuberculosis strains isolated from Finnish patients
    Unpublished
REFERENCE
    2 (bases 1 to 2223)
    Marttila,H.J., Soini,H., Huovinen,P. and Viljanen,M.K.
    Direct Submission
    Submitted (27-NOV-1995) Harri J. Marttila, Mycobacterial Reference
    Laboratory, National Public Health Institute, Kiinamyllykatu 13,
    Turku, SF-20520, Finland
FEATURES
    source
        1. .2223
        /organism="Mycobacterium tuberculosis"
        /strain="H0359/91"
        /db_xref="taxon:1773"
        /note="INH-resistant strain"
    gene
        1. .885
        /gene="katG"
    CDS
        1. .885
        /gene="katG"
        /note="mutant katG"
        /codon_start=1
        /product="catalase-peroxidase"
        /protein_id="AAA85171.1"
        /db_xref="GI:1150704"
        /translation="MPQHPPIETITGASNCPPVGHMKYPVEGGNQDWWPNRLN
        LKVHQNPADVPADAFDAAEVATIDVLTDRIEVMTTSQPNWADYGHYGLP
        IRMAWHAAYRTHDGGGAGGQFAPLNSPNDKARRLLWPPADYGHYGLP
        WADLVFAGNCALESMEGKTFGFGVQDEPVEYWGKEATWLDGERSYGRDLN
        PLAAVMGLIYVNEPQGNPDMAAVDIREFRMAMNDVETALIVGHTFGKTH
        GAGPADLVGPEPEAPLE"

```

```

BASE COUNT 428 a 695 c 740 g 360 t
ORIGIN
Query Match 20.0%; Score 298.4; DB 1; Length 2223;
Best Local Similarity 62.4%; Pred. No. 1.8e-77;
Matches 502; Conservative 0; Mismatches 296; Indels 6; Gaps 2;
QY 688 ATCCCTGGGGGCTGATTTGATTATGCACACAGATTTCAACAGCTGGATATGGAGGTC 747
Db 167 ACCCGATGGTGGGGCTTGCAGCTATCGCGGAGTCCGACCATCGACGTTGACGCC 226
QY 748 TGAATAAGATATCAAGATTTGCTGACAACTTCCAGAGATTGGTCCCTCGGGATTATG 807
Db 227 TGACGCGGACATCGAGGAAGTGTATGACCACTTCGACCGCTGTGGCCGCGACTAG 286
QY 808 GTCAATTATGCTCTTCTTTATTTATGCTTGTGCGGTTCGACGTCGGAACATACAGACAT 867
Db 287 GCCACTACGGCGCTGTTTATCCGATGGCGTGCACGCTCCGCGACCTACCCGATCC 346
QY 868 ATGATGGCGGGAGGCGCCAGTGTGTGTCAGCAACGTTTTTGAACCGCTGAACAGCTGC 927
Db 347 ACGAGCGCGCGCGCGCGCGCGCGCATCGACGGTTCGCGCGCTTAACAGCTGCG 406
QY 928 CGGATAACGTTAATCTGGATAAGCCGCTCGATTGCTGTGGCCAGTCAGAAAAATACG 987
Db 407 CCGACAACGCCAGCTTGGACAAGCGCGCGCGCTGCTGTGGCGCGTCAAGAAGAACTAG 466
QY 988 GCTCCAGTATTTCTGGGAGACCTGATGCTCTGCTGCTGATGTTGCTTGAATCCA 1047
Db 467 GCAAGAGCTCTCATGGCGGACCTGATGTTTTCGCGCACTTCGCGCTGGATCGA 526
QY 1048 TGGGATTTAAACGCTGGGATTTGTGGCGGAAGAGAAGATGACTGGGAGTCGACCTGG 1107
Db 527 TGGGCTTCAAGACGTTTCGGGTTTCGGCTTCGCGCGGTCCACCACTGGGAGCGCGATGAG 586
QY 1108 TATACTGGGGGCTGCACAAACGCTCTGTCAGATAACGGGATA---AAAACGGGAAC 1164
Db 587 TCTATTGGGGCAAGGAAGCCACTGCTGCGGATGAGCGTTACAGCGTTAAGCGGATC 646
QY 1165 TTCAGAAACCTCTGCGCCACGACAGATGGGACTTATTATGCTCAATCTGTAAGCCCG 1224
Db 647 TGGAGAACCCGCTGGCGCGGTGCAGATGGGCTGATCTACGTGAACCCGGAGGCCGA 706
QY 1225 GTGAAACACAGATCTCTGCGTTCGCGGAAGAGATATCAGGAAGCTTTTTCACGTATGG 1284
Db 707 ACGGCAACCCGACCCCATGCGCGCGCGGTGCACATTCGCGAGAGCTTTCGCGCATGG 766
QY 1285 CCATGATGATGAGGAGACTGTGGCCCTGATCGCGGAGGCGATACATTTGGTAAAGCAC 1344
Db 767 CCATGAACACGCTGCAACAGACGCGCGCTGATGCTGCGCGTCCACTTTCGTAAGACCC 826
QY 1345 ATGGTGCAGCGTCTCTGAAATATGTTATGGCGCAGGGCGCTGATGTCACCTGTGGAGG 1404
Db 827 ATGGCGCG--GCCCGCGGATCTGTGCGCGCGCGCAACCCGAGGCTGCTCCGCTGGAGT 883
QY 1405 AGCAGGAGCTGGATGGAATAATATGTTGTCAGGAACCGCAATATACCATCACCA 1464
Db 884 AGATGGCTTGGGCTGGAAGAGCTCGTATGGCACCGGAACCGTAAGGAGCGCATACCA 943
QY 1465 GTGGCTTGAAGAGCGCTGTGCGA 1488
Db 944 CGGCATCGAGTCTGTATGACGA 967

```

RESULT 40
MTU41306

LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE

MTU41306 2223 bp DNA linear BCT 07-JAN-1996
Mycobacterium tuberculosis INH-resistant strain H0055/92,
catalase-peroxidase (katG) gene, complete cds.
U41306
U41306.1 GI:1150705
Mycobacterium tuberculosis strain-H0055/92.

[illegible]

/protein_id="AA85176.1"
/db_xref="GI:1150714"
/translation="MPQHPPITTTTGAASNGCPVVGHMKYFVEGGNODWPNRLN
LKVHONPAVADPMCAAFDIAAEVATIDVDALTRDIEVMTTSQPPWADYGHGPLE
IRMAHAAQTYRIHDCRGAGGGMORFAPLNSWPDNASLDKARLLWPKKYKKLS
WADLVFAGNCALESMEGKTFGFGFRVDQWEDVYWGKEATWLGDERYSKGRDLEN
PLAAYOMGLIYVNPENPNDPMAAAVDIRETFRMANNDVETALIVGGHTFGKTH
GAGPADLVGPEPEAAPELQMGKLGWKSIGTGKDAITSGIEVWNTPTKWDNSFLE
ILYGEWELTKSPAGAWOYTKADGAGAGTIPDFGGRSPKTMATDLSLRVDPIYER
ITRWLHPEELADEFAKAWYKLIHRDMGPVARYLGPLVKOTLLWQDPVPVASHLV
GEADISLKRVTLEIEQESNSAAGNPKVSVFADLVVLGCAALEAKAAGHNITVP
DPEGRDASOEOTDYESFAVLEPKADGFRNLYKGNPLPAEYMLLKANILLISAPEN
FTVGLGRVIGCANYKRLPLGVFTFEASESLTDFVNLDMGITWEPSPADGTYGQKD
GSGKVKVTKGSRVDLVFGNSLRLVVEYGADDAQPKFQVDFVAANDKVMNLDREDFVR

BASE COUNT	428 a	696 c	741 g	358 t
Query Match	20.0%;	Score 298.4;	DB 1;	Length 2223;
Best Local Similarity	62.4%;	Pred. No. 1.8e-77;		
Mismatches	502;	Conservative	0;	Mismatches 296; Indels 6; Gaps 2;
QY	688	ATCCCTGGGGGCTGATTTGATTATGCCACACAGATTTCACACAGCTGGATATGAGGCTC	747	
DB	167	ACCGATGGGTGCGCGCTTCGACTATGCGCGGAGGTCCGACCATCGAGCTTGACGCC	226	
QY	748	TGAAAAAGATATCAAGATTGCTGACAACTTCCAGAGTATGGTCCCTTCGCGGATATG	807	
DB	227	TGACGGGGACATCGAGAAAGTATGACACCTCGCAGCGCTGGTGGCGCGGACTACG	286	
QY	808	GTCATATGTCCTTCTTATTCGTATGCTTGGCAGCGTGGCGGACATACAGACAT	867	
DB	287	GCCACTAGGCGCGCTGTTATCGGATGGGTGGCAGCTGCGGCACCTACCGATCC	346	
QY	868	ATGATGCGCGGAGGCGCAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT	927	
DB	347	ACGACG	406	
QY	928	CGGATACGTTATCTGGATAAGCCCGTGCATGCTGTGGCCAGTCAAGAAAAATACG	987	
DB	407	CCGACACCGCAGCTTGGACAAAGCGCGCGCGCTGCTGTGGCGCGTCAAGAAAGTACG	466	
QY	988	GCTCCAGTATTTCTGCGGAGACCTGATGCTCTGCTGCTGCTGCTGCTGCTGCTGCTG	1047	
DB	467	GCAAGAGCTCTCAATGGCGGACCTGATGTTTTCGCGGCACTGCGCGCTGGAATCGA	526	
QY	1048	TGGGATTTAAACGCTGGGATTTGCTGGCGGAAGAAAGATGACTGGAGTCCGAGCTGG	1107	
DB	527	TGGGCTTCAAGAGCTTCGGGTTTCGGCTTCGCGCGGTTCGACCATCGGAGCCGATGAGG	586	
QY	1108	TATAGTGGGCGCTGACAAACGCTCTTGCAGATAACCGGGATA---AAAACGGAAAC	1164	
DB	587	TCTATTGGGGAAGAAAGCCACCTGGCTCGCGCATGAGCGTTACGCGGTAAACGGGATC	646	
QY	1165	TTCAAGAACCTTCGCGCCACCGATGGACTTATATGTCATCTCTGAAGGCCCGG	1224	
DB	647	TGGAGAACCGCTGCGCGCGGTGCAGATGGGCTGATCTACGTGAACCGCGGAGGCCGA	706	
QY	1225	GTGAAACACGATCTCTGCTGCTGCGCGGAAGAAAGATATCAGGAAGCTTTTTCACGTATGG	1284	
DB	707	ACGCAACCGGACCCCATGCGCGCGGTGCGATTCGACATTCGCGAGAGCTTTCGCGCATGG	766	
QY	1285	CCATGGATGATGAGGAGACTTGGCCCTGATCGCGGAGGCGCATACATTTGTTAAGCAC	1344	
DB	767	CCATGAACGAGCTGCAACAGCGCGCTGATCTCGCGGCTCACACTTCGTTAAGACCC	826	
QY	1345	ATGTCGAGCGCTCTCTGAAATATGTTGGCCAGGCGCTGATGTGTCACCTGTGGAG	1404	
DB	827	ATGCGCGG---GCCGCGCGGATCTGGTGGCGCGCGGACCCGAGGCTGCTCGCTGGAGC	883	
QY	1405	AGCAGGACTGGATGAAATAAATGTGGTACAGGAAACCGCAATATACCATCACCA	1464	

884 AGATGGGCTGGCTGGAAGAGCTCGTATGGCACCGGAACCGTAAGACCGCATACCA 943
1465 GTGGCTGGAAGAGCTCGTCTGCA 1488
944 GCGCATCGAGGCTCGTATGACCA 967
MTU41313 2223 bp DNA linear BCT 07-JAN-1996
Mycobacterium tuberculosis INH-resistant strain H0211/94,
catalase-peroxidase (katG) gene, complete cds.
U41313
U41313.1 GI:1150719
Mycobacterium tuberculosis strain-H0211/94.
Mycobacterium tuberculosis
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
Actinomycetales; Corynebacterineae; Mycobacteriaceae;
Mycobacterium; Mycobacterium tuberculosis complex.
1 (bases 1 to 2223)
katG gene mutations in Isoniazid-resistant Mycobacterium
Martiila,H.J., Soini,H., Huovinen,P. and Viljanen,M.K.
tuberculosis strains isolated from Finnish patients
Unpublished
2 (bases 1 to 2223)
Martiila,H.J., Soini,H., Huovinen,P. and Viljanen,M.K.
Submitted (27-NOV-1995) Harri J. Martiila, Mycobacterial Reference
Laboratory, National Public Health Institute, Klimamyllynkatu 13,
Turku, SF-20520, Finland
Location/Qualifiers
1..2223
/organism="Mycobacterium tuberculosis"
/strain="H0211/94"
/db_xref="taxon:1773"
/note="INH-resistant strain"
1..2223
/gene="katG"
1..2223
/gene="katG"
/codon_start=1
/transl_table=11
/product="catalase-peroxidase"
/protein_id="AA85179.1"
/db_xref="GI:1150720"

BASE COUNT	428 a	695 c	741 g	359 t
Query Match	20.0%;	Score 298.4;	DB 1;	Length 2223;
Best Local Similarity	62.4%;	Pred. No. 1.8e-77;		
Mismatches	502;	Conservative	0;	Mismatches 296; Indels 6; Gaps 2;
QY	688	ATCCCTGGGGGCTGATTTGATTATGCCACAGATTTCACACAGCTGGATATGAGGCTC	747	
DB	167	ACCGATGGGTGCGCGCTTCGACTATGCGCGGAGGTCCGACCATCGAGCTTGACGCC	226	
QY	748	TGAAAAAGATATCAAGATTTCGTCAGAACTCCACAGATTGGTCCCTCGCGATTATG	807	

Db 647 TGGAGAACCCCGCTGGCCGCGTGCAGATGGGCTGATCTACGTGAACCCGAGGGGCCGA 706
 QY 1225 GTGGAACACAGATCTCTGGCTTCGCGAAAGATATCAGGAAGCTTTTTCACGTATGG 1284
 Db 707 ACGGACACCCGAGCCCATGCGCGGCGGTGCGACATTCGGGACAGCTTCGGCGCATGG 766
 QY 1285 CCATGGATGATGAGGAGACTGTGCCCTGTGCGGGAGGCGCATPACATTTGGTAAGCAC 1344
 Db 767 CCATGAACGAGCTGGAACACGCGGCTGATCGTCGGCGGTTCACACATTCGTTAAGACCC 826
 QY 1345 ATGGTGAGCTCTCTGTAATAATGATTTGGCGCAGGCGCTGATGTCACCTGTGGAGG 1404
 Db 827 ATGGCGCG---GCCCGCCGATCTGGTGGCCCGCGAACCAGGCTGCTCCGCTGGAGC 883
 QY 1405 AGCAGGAGCTGGGATGGAATAAATGATGTCAGGAAACGCGCAATATATACCATACCA 1464
 Db 884 AGATGGCTTGGCTGGAGAGCTGCTATGGCACCGGACCGGTAAGGACCGCATACCA 943
 QY 1465 GTGGCTGGAAGGAGCTGTGTCGA 1488
 Db 944 GCGGCATCGAGTCTGATGGACGA 967

RESULT 44

MTU41304
 LOCUS
 DEFINITION
 Mycobacterium tuberculosis H37Rv, BCT 07-JAN-1996
 catalase-peroxidase (katG) gene, complete cds.
 U41304
 U41304.1 GI:1150701

ORGANISM
 Mycobacterium tuberculosis strain-H0078/91.

REFERENCE
 AUTHORS
 Marttila, H.J., Soini, H., Huovinen, P. and Viljanen, M.K.
 TITLE
 katG gene mutations in isoniazid-resistant Mycobacterium tuberculosis strains isolated from Finnish patients
 JOURNAL
 Unpublished

REFERENCE
 AUTHORS
 Marttila, H.J., Soini, H., Huovinen, P. and Viljanen, M.K.
 TITLE
 Direct Submission
 Submitted (27-NOV-1995) Harri J. Marttila, Mycobacterial Reference Laboratory, National Public Health Institute, Kilinmallynkatu 13, Turku, SF-20520, Finland
 JOURNAL
 Location/Qualifiers

FEATURES
 source
 1..2225
 /organism="Mycobacterium tuberculosis"
 /strain="H0078/91"
 /db_xref="taxon:1773"
 /note="INH-resistant strain"

gene
 1..1875
 /gene="katG"
 /note="mutant katG"

CDS
 1..1875
 /codon_start=1
 /transl_table=11
 /product="catalase-peroxidase"

protein_id="AA085170.1"
 /db_xref="GI:1150702"

translation="MPEQHPPITTTTGAASNGCPVVGKMKYPVEGGNODWPNRLN
 LKVLHQNFAVDPMGAADYAEVATIDVATLTDIEVMTTSQPMFADYGHYGLP
 IRMAWEAAGTVRIHDGRGAGGGMORFAPLNSWPDNSLDKARLLMPVKYKKLS
 WADLVFAGNCALESMSGFKTFGFGVDQPDVYWGKATWLDGDRYSGKRDLEN
 PLAAVQMGILVYNPSPNGNDPMAAADVIRETFRMAWNVETAAALVGGHTGKTH
 GAGPADLVGPEPEAPLEQMLGHWKSSVTGTGDKAITSGLVETWNTPTKWDNSFLE
 ILYGYEWELTSPAGAWYTAQDGAGTIPDPFGPGRSPTMLATDLSLRVDPPIER
 ITRRWLEHPEELDAFAKAWYKLIHRDMPVARYLGPLVPKOTLWDPVPAVSHDLV
 GAEATLSKRSQASGLTASOLVSTAWAAASFRGSKRGANGRRIRLOPQVGWEVN
 DFDGDLRKVIRLTLEIESFNSAPGNIKVSFADLVVLGGCAAITEKAAKAAGHNTVP

FTPGRDASQEQDVFESFAVLEPKADGFRNYLGNPLPAETWMLDKANHLRSVPLR
 BASE COUNT 428 a 695 c 741 g 359 t. 2 others
 ORIGIN
 Query Match 20.0%; Score 298.4; DB 1; Length 2225;
 Best Local Similarity 62.4%; Pred. No. 1.8e-77;
 Matches 502; Conservative 0; Mismatches 296; Indels 6; Gaps. 2;
 QY 688 ATCCCTGGGGGCTGATTTTATGATGACCAAGATTTCAACAGCTGGATATGAGGCTC 747
 Db 167 ACCGATGGTGGCGGTTCGACTATGCCGGAGGTTCGGACCATCGAGTTGACGCC 226
 QY 748 TGAATAAGATATCAAGATTTCTGACAACTTCCAGAGTGTGTCCTCGGATATG 807
 Db 227 TGACGGGACATCGAGGAAGTATGACCACTCGCAGCGGTGTGCGCCGCGACTACG 286
 QY 808 GTCAATATGCTCTCTTTTATTCGATGCTTGGCAGCGTGGCGGACATACAGACAT 867
 Db 287 GCCACTACGGCGCGCTGTTTATCCGATGCGTGGCAGCTGCGGCACCTACCGCATCC 346
 QY 868 ATGATGGCGGGGAGCGCCAGTGGTGCAGCAAGCTTTTGAACCGCTGAACAGTGGC 927
 Db 347 ACNAGCGCGCGCGCGCGCGCGCGCATGACGGGTTCGCGCGCTTAACAGCTGGC 406
 QY 928 CGGATAACGTTAATCTGGATAAAGCGCGTGTGTCGTCGCGCAGTCAAGAAAATACG 987
 Db 407 CCGACAACGCGAGCTTGGACAAGCGCGCGGTGTGTCGCGCGGTCAAGAAAGTACG 466
 QY 988 GCTCCAGTATTTCTGGGAGACCTGATGCTCTGCTGCTGATGTTGTCCTTGAATCCA 1047
 Db 467 GCAAGAAGCTCTCATGGCGGACCTGATTTTTCGCGCGCACTGCGCGCTGGAATCGA 526
 QY 1048 TGGGATTTAAACGCTGGGATTTTCTGGCGGAAGAGAGATGACTGGAGTCCGACCTGG 1107
 Db 527 TGGGCTTCAAGAGCTTGGCTTCGGCTTCGGCTTCGGCGGTGCGACAGTGGGAGCCGATGAGG 586
 QY 1108 TATCTGGGCGCTGACAAACGCTTTCGACATACCGGGATA---AAAACGGGAAC 1164
 Db 587 TCTATTGGGGCAAGGAACCCCTGGCTCGCGCATGAGCGTTACAGCGGTAAGCGGGATC 646
 QY 1165 TTCAGAAACCTCTTGGCGCCAGCAGATGGGACTTATTTATGTCAATCTCTGAAGGCCCG 1224
 Db 647 TGGAGAACCCGCTGGCGCGGTGCAGATGGGCTGATCTACGTGAACCCGAGGGCCGA 706
 QY 1225 GTGGAACACAGATCTCTGCTGCTCCGCGAAGATATCAGGAGCTTTTTCACGTATGG 1284
 Db 707 ACGCAACCGGACCCCATGGCGCGGTTCGACATTCGCGAGACGTTTTCGGCGCATGG 766
 QY 1285 CCATGGATGATGAGGAGACTGTGGCCCTGATCGCGGAGGCGCATACATTTGGTAAAGCAC 1344
 Db 767 CCATGAACGAGCTCGAAACAGCGCGCTGATCGTGGCGGTTCACATTTTCGGTAAAGACC 826
 QY 1345 ATGGTGAGGCTCTCTCTGAATAATGATTTGGCGCAGGCGCTGATGTCACCTGTGGAGG 1404
 Db 827 ATGGCGCG---GCCCGCGCATCTGCTCGCGCCCGAACCAGGCTGCTCGCTGGAGC 883
 QY 1405 AGCAGGAGCTGGGATGGAATAAATATGTCGTCAGGAAACGCGCAATATATACCATACCA 1464
 Db 884 AGATGGCTTGGCTGGAGAGCTGCTATGGCACCGGACCGGTAAGGACCGCATACCA 943
 QY 1465 GTGGCTGGAAGGAGCTGTGTCGA 1488
 Db 944 GCGGCATCGAGTCTGATGGACGA 967

RESULT 45

LOCUS
 I61339
 DEFINITION
 Sequence 1 from patent US 5658733.
 ACCESSION
 I61339
 VERSION
 I61339.1
 KEYWORDS
 I61339 2235 bp DNA linear PAT 07-OCT-1997

RESULT 46	LOCUS	I74293	Sequence 1 from patent US 5688639.	DNA	linear	PAT 03-APR-1998
DEFINITION	ACCESSION	I74293	I74293			
KEYWORDS	VERSION	I74293.1	GI:3010434			
SOURCE	ORGANISM	Unknown.	Unknown.			
REFERENCE	1 (bases 1 to 2235)					
AUTHORS	Cockerill, F.R., Kline, B.C. and Uhl, J.R.					
TITLE	Detection of isoniazid resistant strains of M. tuberculosis					
JOURNAL	Patent: US 5688639-A 1 18-NOV-1997;					
FEATURES	Location/Qualifiers					
source	1..2235					
BASE COUNT	431 a 696 c 745 g 363 t					
ORIGIN						
Query Match	20.0%;	Score 298.4;	DB 6;	Length 2235;		
Best Local Similarity	62.4%;	Pred. No. 1.8e-77;				
Matches	502;	Conservative	0;	Mismatches 296;	Indels	6;
Gaps	2;					
QY	688	ATCCCTGGGGGCTGATTTTGTATATGTCACACAGATTTCAACAGCTGGATATGAGGCTC	747			
DB	176	ACCCGATGGTGGGGCTGACATATGCGCGGAGTGGCGACCATCGAGCTTTCAGCGCC	235			
QY	748	TGAAAAAGATATCAAGATTTGCTGACAACTTCCAGAGTGTGGTCCCTGCGGATATG	807			
DB	236	TGACCGGGACATCAGGAAGTATGACCACTCGACCGCTGTGGCGCCGACTACG	295			
QY	808	GTCAATATGCTCTTCTTTATTCGTATGGCTGGCACAGTTCGCGGAGTGGCGACATACAGGACAT	867			
DB	296	GCCACTACGGGCGCTGTTATCCGATGGGTGGCGACGCTGCGGACCATCCGCATCC	355			
QY	868	ATGATGGCGGGAGCGCCAGTGTGTGTCAGCAAGCTTTTGAACCCCTGAACAGCTGGC	927			
DB	356	ACGACGGCGCGCGCGCGCGGCGCATGCGAGCGGTTCGCGCGCTTAACAGCTGGC	415			
QY	928	CGGATAAGCTTAATCTGGATAAAGCCCTCGATTTGCTGTGCCAGTCAAGAAAAAATACG	987			
DB	416	CGACAAACGCCAGCTTGACAAAGCGCGCGCTGTGCGCGGTCAAGAGAGTACG	475			
QY	988	GCTCCAGTATTTCTGGGAGACCTGTGCTGACGTGTAATGTTGCCCTTCAATCCA	1047			
DB	476	GCAAGAGCTCATGGGCGACCTGATTTTCCGCGGCACTGCGCGTGGATCGA	535			
QY	1048	TGGGATTTAAACGCTTGGATTTGCTGGCGAAGAGAAGTACTGTGGAGTTCGAGCTGG	1107			
DB	536	TGGGCTTCAAGAGCTTGGGCTTCGGCTTCGGCGGCTGACCTGAGCGGTAAAGCGGATC	595			
QY	1108	TATACGTGGGCTGACAAACGCTTTCGAGATAACCGGATA--AAACGGGAAC	1164			
DB	596	TCATTGGGCAAGGACCACTGGCTGCGGAGTGGCTGATCTACGTAAACCGGAGGATC	655			
QY	1165	TTGAGAACTCTTCGCGGAGTGGGCTGATTTATGCTCAATCTGAAGGCGCCG	1224			
DB	656	TGGGAAACCGCTGCGCGGCTGAGATTTGGGCTGATCTGCGGAGACGTTTCGGCGCATG	715			
QY	1225	GTGGAACCAAGATCTCTGGCTTCGCGGAAAGATATACAGGAAGCTTTTTCAGGTATG	1284			
DB	716	ACGGCAACCGGACCCATGGCGGCGGCTGACATTCGGGAGACCTTCGGCGCATG	775			
QY	1285	CCATGATGATGAGGAGCTGTGGCCCTGATCGCGGAGGACATATTTGTAAGACAC	1344			
DB	776	CCATGACGAGCTCGAAACAGCGGCTGATCGTGGCGGCTGACACTTCGGTAAGACCC	835			
QY	1345	ATGGTCAGCGCTCTCTGAAATGATTTGGCGAGGCGCTGATGGTGCACCTGTGGAGG	1404			
DB	836	ATGGCGCG--GCCGCGGATCTGGTCGGCGGCAACCGAGGCTGCTCCGCTGGAGC	892			
QY	1405	ACGAGGATGGATGGAATAAATGTTGTACAGGAACCGGAATATACCATACCA	1464			
DB	893	AGATGGGCTTGGCTGGAAGAGCTGATGGACCGGACCGGTAAGGACGCGATACCA	952			
QY	1465	GTGGCTGGAAGAGCGCTGCTGA 1488				
DB	953	CGGGATCGAGGCTGATGAGCA 976				

QY	1405	ACGAGGACTGGGATGGAAAAATAAATGTGGTACAGGAAACGCCAAATATACATCACC	1464
Db	893	AGATGGGCTTGGGCTGGAGAGAGCTGTATGGTACCGGAACCGGTAAAGGACGGCATCACC	952
QY	1465	GTGGCCTGGAAGGAGCGCTGGTCGA	1488
Db	953	GCGGCATCGAGGTCGTATGGACCA	976

RESULT	47
MTU06269	
LOCUS	
DEFINITION	MTU06269 Mycobacterium tuberculosis I33308 catalase (katG) gene, complete cds.
	linear BCT 20-MAY-1994

U06269
 VERSION
 U06269.1 GI:488453
 KEYWORDS
 SOURCE
 ORGANISM
 Mycobacterium tuberculosis.
 Mycobacterium tuberculosis.
 Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium
 tuberculosis complex

REFERENCE
AUTHORS
TITLE
1 (sites)
Cockerill, F. R., Uhl, J. R., Tenesnes, Z., Zhang, Y., Stockman, L.,
Roberts, G. D., Williams, D. L. and Kline, B. C.
Rapid identification of a point mutation of the *Mycobacterium*
tuberculosis catalase-peroxidase (*katG*) gene associated with
isoniazid resistance

JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2297)
AUTHORS Kline, B.C.

TITLE Direct Submission
JOURNAL Submitted (07-FEB-1994) Bruce C. Kline, Mayo Clinic and Mayo Foundation, Biochemistry and Molecular Biology, 200 SW 1 St., Rochester, MN 55905, USA

```

FEATURES             Location/Qualifiers
     1..2297
     /organism="Mycobacterium tuberculosis"
     /strain="I33308"
     /db_xref="taxon:1773"
     70..2292
     /gene="katG"
     70..2292
     /gene="katG"
     /standard_name="catalase-peroxidase"
     /codon_start=1
     /transl_table=11
     /product="catalase"
     /protein_id="AA18239.1"
     /db_xref="GI:488454"

```

[illegible]

BASE COUNT	447 a	718 c	754 g	373 t	5 others
ORIGIN					
Query Match		20.0%	Score 298.4;	DB 1;	Length 2997;
Best Local Similarity		62.4%;	Pred. No. 1.8e-77;		
Matches 502;	Conservative	0;	Mismatches 296;	Indels	6; Gaps 2;
QY	688	ATCCCTGGGGGGCTGATTTTGATTATCCACCACGATTTCAACAGCTGGATATGGAGGCTC	747		

Db	236	ACCGATGGGTGGCGGCTTCGACTATCCCGCGAGGTGCGACCATCGAGCTTGACGCC	295
Qy	748	TGAAAAAGATATCAAGATTTGCTGACAACTTCCAGGATTTGGTCCCTGCGGATTATG	807
Db	296	TGACGGGACATCGAGGAAGTGATGACCACCTCGACCGCTGGTGGCCGCGGACTACG	355
Qy	808	GTCAATTAAGTCTCTTTATTTCGTATGCTTTGGCACGGTGGCGGAACATACAGACAT	867
Db	356	GCCACTACGGCGCGCTTTTATCCGATGCGTGGCACCGTGTGCGGCACCTACCGCATCC	415
Qy	868	ATGATGCGCGGAGGCGCCAGTGGTGTCAGCAACGTTTTGAACCGCTGAACAGCTGGC	927
Db	416	ACGACGCGCGCGCGCGCGGGCGGATGACGCGGTTGCGCGCGCTTAAACAGCTGGC	475
Qy	928	CGGATAACGTTAATCTCGATTAAGCCCGTCGATTGCTGTGCCCAGTCAAGAAAAATACG	987
Db	476	CCGACAAGCCAGCTTGGACAAGCGCGCGGCTGCTGTGCCGGTCAAGAAGAATACG	535
Qy	988	GCTCCAGTATTTCTGGGGAGACCTGATGCTCTGACTGTGTATGTTGCCCTTGAATCCA	1047
Db	536	GCAAGAAGCTCTCATGSGGAGACCTGATTGTTTTGCGCGCAACTGCGCGCTGGAATCGA	595
Qy	1048	TGGGATTTAAACGCTGGGATTTGCTGCGCGAAGAGAAGATGACTGGGAGTTCGACCTGG	1107
Db	596	TGGGCTTCAAGACGTTGCGGTTGCGGCTTCGCGCGGGTCGACCATGGGAGCCCGATGAGG	655
Qy	1108	TATACTGGGGCGCTGACAAACGCTCTTCAGATAACCGGGATA--AAAACGGGAAC	1164
Db	656	TCTATTGGGCAAGGAAGCACCTGGCTCGCGCATGAGCGTTACAGCGGTAAAGCGGATC	715
Qy	1165	TTCAGAAACCTCTGGCGCCACGAGATGGACTTATTATGTCACTCTGAAAGGCCCG	1224
Db	716	TGGAGAACCCTGGCCGCGGTCAGATGGGGCTGATCTACGTGAACCGCGAGGGGCGGA	775
Qy	1225	GTGAAAAACAGATCCTCTGGCTTCCCGCAAGAGATATCAGGGAAGCTTTTTCACGTATGG	1284
Db	776	ACGCAACCGCGACCCCATGGCCGCGCGTTCGACATTCGCGAGACGTTTCGGCGCATGG	835
Qy	1285	CCATPGATGATGAGGAGACTGTGCCCTGTATCGCGGGAGGCGATACATTTGGTAAAGCAC	1344
Db	836	CCATGAACAGCTCGAAACACGCGCGCTGATCTGCTCGCGCGTCACTTTTCGGTAAAGACC	895
Qy	1345	ATGGTGCAGGCTCTCCTGAAATGTTATGGCGAGGCGCTGTGTTGTCACCTGTGGAGG	1404
Db	896	ATGGCGCG--GCCCCGCGATCTGTGTCGCCCCCGAACCAGGCTGTCTCCGCTGGAGC	952
Qy	1405	ACGAGGACTGGATGAAAAATAAATGTGTTACAGGAACCGCAATATACCATCACCA	1464
Db	953	AGATGGCTTGGGCTGGAAGAGCTCGTATGGCACCGGAACCGGTAAAGACGCGTACCA	1012
Qy	1465	GTGCCCTGGAGAGCCTGTGTCGA	1488
Db	1013	CGGGCATCGAGGTCGTATGGACGA	1036

RESULT 48	2311 bp	DNA	linear	BCT 23-JUL-1996
LOCUS	MTU06258			
DEFINITION	Mycobacterium tuberculosis ATCC 25618 catalase (katG) gene, complete cds.			
ACCESSION	U06258			
VERSION	U06258.1	GI:1439530		
KEYWORDS	Mycobacterium tuberculosis.			
SOURCE	Mycobacterium tuberculosis			
ORGANISM	Bacteria; Firmicutes; Actinobacteria; Actinobacteridae; Actinomycetales; Corynebacteriaceae; Mycobacteriaceae; Mycobacterium; Mycobacterium tuberculosis complex.			
REFERENCE	1 (sites)			
AUTHORS	Cockerill, F.R., Uhl, J.R., Temesgen, Z., Zhang, Y., Stockman, L., Roberts, G.D., Williams, D.L. and Kline, B.C.			
TITLE	Rapid identification of a point mutation of the Mycobacterium tuberculosis catalase-peroxidase (katG) gene associated with			

1108	TATACTGGGGGCGCTTGCACAAAGACCTCTTGCAGATAAACCGGGATA---	1166	AAAAACGGGAAAC
636	TCTATTGGGGCAAGCAAGCCACCTGGCTCGCGATGAGCGTTACAGCGGTAAAGCGGATC	695	
1165	TTCAGAAACCTCTTCCGCCACGAGATGGGACATTTATTTGTCATCTTCAAGGCGCCG	1224	
696	TGAGAACCGCTGGCGCGGTGCAGATGGGGCTGATCTAGCTGAACCCGAGGGGCGCA	755	
1225	GTGGAAACACAGATCTCTGCTTCCGCGAAAGATATCAGGGAAGCTTTTTCAGGTATG	1284	
756	ACGGCAACCCGACCCCATGGCGGGCGGTGACATTCGGGACAGCTTTCGGCGCATGG	815	
1285	CCATGGATGATGAGGAGACTGTGGCCCTGATCGCGGGAGGACATATTTGGTAAAGCAC	1344	
816	CCATGAACGACGTCGAAACAGCGGCGCTGATCGTGGCGGTTCACACTTTCGGTAAGACC	875	
1345	ATGGTGCAGCTCTCTGNAANAATGATTTGGCGAGGCGCTGTGGTGCACCTGTGGAG	1404	
876	ATGGCGCGC---GCCCGCGCATCTGTCGCGCCCGAAACCCGAGGCTCTCCGCTGGAGC	933	
1405	ACGAGGACTGGGATGGAATAAATGTTGTACAGGAAACCGCAATATACCATCACCA	1464	
933	AGATGGCTTGGGCTGGAGAGCTCGTATGGCAGCAACCGTAAGGACCGATACCA	992	
1465	GTGGCTGGAAGGACCTGGTCTGA 1488		
993	CGCGCATCGAGCTCGTATGAGCA 1016		
RESULT 49			
MTU06264			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
FEATURES			
source			
gene			
CDS			

1108	TATACTGGGGGCGCTTGCACAAAGACCTCTTGCAGATAAACCGGGATA---	1166	AAAAACGGGAAAC
636	TCTATTGGGGCAAGCAAGCCACCTGGCTCGCGATGAGCGTTACAGCGGTAAAGCGGATC	695	
1165	TTCAGAAACCTCTTCCGCCACGAGATGGGACATTTATTTGTCATCTTCAAGGCGCCG	1224	
696	TGAGAACCGCTGGCGCGGTGCAGATGGGGCTGATCTAGCTGAACCCGAGGGGCGCA	755	
1225	GTGGAAACACAGATCTCTGCTCCGCGAAAGATATCAGGGAAGCTTTTTCAGGTATGG	1284	
756	ACGGCAACCCGACCCCATGGCGGGCGGTGACATTCGCGAGACGTTTCGGCGCATGG	815	
1285	CCATGGATGATGAGGAGACTGTGGCCCTGATCGCGGGAGGACATATTTGGTAAAGCAC	1344	
816	CCATGAACGACGTCGAAACAGCGGCGCTGATCGTGGCGGTTCACACTTTTCGGTAAGACC	875	
1345	ATGGTGCAGCTCTCTGNAANAATGATTTGGCGAGGCGCTGATGGTGCACCTGTGGAGG	1404	
876	ATGGCGCGC---GCCCGCGCATCTGTCGCGCCCGAAACCCGAGGCTCTCCGCTGGAGC	933	
1405	ACGAGGACTGGGATGGAATAAATGTTGTACAGGAAACCGCAATATACCATCACCA	1464	
933	AGATGGCTTGGGCTGGAGAGCTCGTATGGCAGCAACCGTAAGGACCGATACCA	992	
1465	GTGGCGTGAAGGACCTGGTGTGA	1488	
993	CGCGCATCGAGCTGATATGAGCA	1016	
RESULT 49			
MTU06264			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
FEATURES			
source			
gene			
CDS			

Db 356 GCACCTACGGCGCGCTTTATCGGNTGCGTGACGCTGCGGGACCTACGCTACC 415
Qy 868 ATGATGCGCGGGAGCGCCAGTGGTGCAGCAACGTTTGAACCGCTGAACAGCTGGC 927
Db 416 ACACGCGCGCGCGCGCGCGCGCATGACGCGGTTCCGCGCGCTTAACAGCTGGC 475
Qy 928 CGGATACGTTAATCTGGTAAGACCGCTCGATGCTGTGGCCAGTCAAGAAAATACG 987
Db 476 CCGACAACACGAGTTGGACAAGCGCGCGCGCTGTGTGGCGCGTCAAGAAGATACG 535
Qy 988 GCTCCAGTATTCTGGGGAGACTGATGCTCTGACTGTAATGTTGCCCTTGAATCCA 1047
Db 536 GCAGAGACTCTATGCGCGGAGCTGATGTTTTCGCCGCACTGCGCGCTGGATCGA 595
Qy 1048 TGGGATTTAAACGCTGGGATTTCTGGCGGGAAGAGATGACTGGGAGTTCGGACCTGG 1107
Db 596 TGGGCTTCAAGACGTTTCGGGTTTCGGGTTTCGGCGGGTCGACCACTGTTGGGAGCCGATGAGG 655
Qy 1108 TATACGGGGGCTGACACAAGCCCTTTCAGATACCGGGATA---AAAACGGGAAC 1164
Db 656 TCATTGGGCAAGAGACACCTGCGCTCGCGGATGAGCGTTACAGCGGTAAGCGGGATC 715
Qy 1165 TTCAGAAACCTCTGCGCCACACAGATGGGACTTATTTATGTCAATCCTGAAGGCCCG 1224
Db 716 TGGAGAACCGCTGGCGCGCGGTGCAGATGGGGTGTACTAGTGAACCCGGAGGGCCGA 775
Qy 1225 GTGGAAACAGATCCTCTCGGCTTCCGCGAAAGATATCAGGGAAGCTTTTTCACGTATGG 1284
Db 776 ACGCAACCGGACCCCATGCGCGCGGTCGACATTCGAGAGACGTTTCGGCGCATGG 835
Qy 1285 CCATGGATGATGAGGAGACTGTGGCCCTGATCGCGGGAGGGCATACATTTGGTAAAGCAC 1344
Db 836 CCATGAACGACGTGGAACACAGCGCGCTGATCGCGCGGTACACACTTTCGGTAAAGACC 895
Qy 1345 ATGTCGACGCTCTCCTGAAAATGTAATGGCGCAGGGCCTGATGGTGACCTGTGGAGG 1404
Db 896 ATGGCGCG---GCCCGCGGATCTGCTCGCGCCCGAACCAGGCTGCTCCGCTGGAGC 952
Qy 1405 AGCAGGACTGGGATGGAATAAATGTTGGTACAGGAACGGCAATAATACCATCACCA 1464
Db 953 AGATGGGCTTGGGCTGGAAGAGCTCGTATGGCACCGGACCGGTAAGGACGCGATCACCA 1012
Qy 1465 GTGGCCTGGAAGGAGCCTGGTCGA 1488
Db 1013 GCGGCATCGAGGTCGTATGGACGA 1036

Search completed: June 18, 2003, 04:52:44
Job time : 2697.55 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 18, 2003, 04:53:09 ; Search time 376.433 Seconds
(without alignments)
5804.490 Million cell updates/sec

Title: US-09-674-277-1
Perfect score: 1489
Sequence: 1 ctgcagtcgagatgaaag.....ctggaagagcctgctgac 1489

Scoring table: IDENTIFY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1042519 seqs, 733713590 residues

Total number of hits satisfying chosen parameters: 2085038

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 150 summaries

Database : Published Applications_NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
- 12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
- 13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
- 14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	357.6	24.0	2238	10	US-09-884-889-7
2	224.6	15.1	2262	10	US-09-884-889-5
3	38.6	2.6	567	9	US-10-123-155-254
4	34.6	2.3	520	9	US-10-184-644-332
5	34.6	2.3	520	9	US-10-184-634-332
6	34.2	2.3	893	10	US-09-815-242-2188
7	34.2	2.3	927	10	US-09-815-242-8989
8	34.2	2.3	1164	10	US-09-815-242-4832
9	34.2	2.3	2354	7	US-08-781-986A-565
10	34	2.3	636	9	US-09-764-891-8918
11	33.8	2.3	457	10	US-09-864-761-3110
12	33.8	2.3	499	9	US-10-184-644-592
13	33.8	2.3	499	9	US-10-184-634-592
14	33.8	2.3	579	9	US-09-924-400-147
15	33.8	2.3	579	10	US-09-810-936-147
16	33.8	2.3	579	10	US-09-429-755-147
17	33.8	2.3	882	9	US-10-184-644-574
18	33.8	2.3	882	9	US-10-184-634-574
19	33.4	2.2	6383	9	US-10-092-154-1415

c 20	33.4	2.2	6383	10	US-09-764-847-1415	Sequence 1415, Ap
c 21	33.4	2.2	21732	9	US-10-092-154-1414	Sequence 1414, Ap
c 22	33.4	2.2	21732	10	US-09-764-847-1414	Sequence 1414, Ap
c 23	33.2	2.2	484	9	US-09-918-995-29551	Sequence 29551, A
24	33.2	2.2	3226	9	US-09-992-598-352	Sequence 352, App
25	33.2	2.2	3226	9	US-09-989-293A-352	Sequence 352, App
26	33.2	2.2	3226	9	US-09-989-733-352	Sequence 352, App
27	33.2	2.2	3226	9	US-09-990-444-352	Sequence 352, App
28	33.2	2.2	3226	9	US-09-989-730-352	Sequence 352, App
29	33.2	2.2	3226	9	US-09-990-436-352	Sequence 352, App
30	33.2	2.2	3226	9	US-09-991-181-352	Sequence 352, App
31	33.2	2.2	3226	9	US-09-993-687-352	Sequence 352, App
32	33.2	2.2	3226	9	US-09-989-734-352	Sequence 352, App
33	33.2	2.2	3226	9	US-10-028-072-463	Sequence 463, App
34	33.2	2.2	3226	9	US-09-997-653-352	Sequence 352, App
35	33.2	2.2	3226	9	US-09-993-667-352	Sequence 352, App
36	33.2	2.2	3226	9	US-10-121-049-463	Sequence 463, App
37	33.2	2.2	3226	9	US-10-123-904-463	Sequence 463, App
38	33.2	2.2	3226	9	US-10-140-470-463	Sequence 463, App
39	33.2	2.2	3226	9	US-09-990-438-352	Sequence 352, App
40	33.2	2.2	3226	9	US-09-990-562-352	Sequence 352, App
41	33.2	2.2	3226	9	US-09-997-428-352	Sequence 352, App
42	33.2	2.2	3226	9	US-09-997-666-352	Sequence 352, App
43	33.2	2.2	3226	9	US-10-175-746-463	Sequence 463, App
44	33.2	2.2	3226	9	US-10-176-918-463	Sequence 463, App
45	33.2	2.2	3226	9	US-10-176-921-463	Sequence 352, App
46	33.2	2.2	3226	9	US-09-990-711-352	Sequence 352, App
47	33.2	2.2	3226	9	US-10-137-865-463	Sequence 463, App
48	33.2	2.2	3226	9	US-10-140-474-463	Sequence 463, App
49	33.2	2.2	3226	9	US-10-142-431-463	Sequence 463, App
50	33.2	2.2	3226	9	US-10-143-114-463	Sequence 463, App
51	33.2	2.2	3226	9	US-09-989-726-352	Sequence 352, App
52	33.2	2.2	3226	9	US-10-140-002-463	Sequence 463, App
53	33.2	2.2	3226	9	US-09-990-437-352	Sequence 352, App
54	33.2	2.2	3226	9	US-09-998-156-352	Sequence 352, App
55	33.2	2.2	3226	9	US-10-142-419-463	Sequence 463, App
56	33.2	2.2	3226	9	US-09-991-157-352	Sequence 352, App
57	33.2	2.2	3226	9	US-09-991-172-352	Sequence 352, App
58	33.2	2.2	3226	9	US-09-997-514-352	Sequence 352, App
59	33.2	2.2	3226	9	US-09-997-573-352	Sequence 352, App
60	33.2	2.2	3226	9	US-10-123-262-463	Sequence 463, App
61	33.2	2.2	3226	9	US-10-142-423-463	Sequence 463, App
62	33.2	2.2	3226	9	US-09-990-443-352	Sequence 352, App
63	33.2	2.2	3226	9	US-09-990-726-352	Sequence 352, App
64	33.2	2.2	3226	9	US-09-997-559-352	Sequence 352, App
65	33.2	2.2	3226	9	US-09-997-601-352	Sequence 352, App
66	33.2	2.2	3226	9	US-10-121-050-463	Sequence 463, App
67	33.2	2.2	3226	9	US-10-141-755-463	Sequence 463, App
68	33.2	2.2	3226	9	US-09-989-729A-352	Sequence 352, App
69	33.2	2.2	3226	9	US-09-990-440-352	Sequence 352, App
70	33.2	2.2	3226	9	US-09-991-854-352	Sequence 352, App
71	33.2	2.2	3226	9	US-09-997-349-352	Sequence 352, App
72	33.2	2.2	3226	9	US-09-997-440-352	Sequence 352, App
73	33.2	2.2	3226	9	US-09-997-628-352	Sequence 352, App
74	33.2	2.2	3226	9	US-09-997-683-352	Sequence 352, App
75	33.2	2.2	3226	9	US-10-143-032-463	Sequence 463, App
76	33.2	2.2	3226	9	US-09-993-469-352	Sequence 352, App
77	33.2	2.2	3226	9	US-09-993-748-352	Sequence 352, App
78	33.2	2.2	3226	9	US-09-997-542-352	Sequence 352, App
79	33.2	2.2	3226	9	US-10-123-108-463	Sequence 463, App
80	33.2	2.2	3226	9	US-10-123-236-463	Sequence 463, App
81	33.2	2.2	3226	9	US-10-123-261-463	Sequence 463, App
82	33.2	2.2	3226	9	US-10-140-921-463	Sequence 463, App
83	33.2	2.2	3226	9	US-10-140-928-463	Sequence 463, App
84	33.2	2.2	3226	9	US-09-990-427-352	Sequence 352, App
85	33.2	2.2	3226	9	US-09-990-439-352	Sequence 352, App
86	33.2	2.2	3226	9	US-10-121-045-463	Sequence 463, App
87	33.2	2.2	3226	9	US-10-123-292-463	Sequence 463, App
88	33.2	2.2	3226	9	US-10-123-903-463	Sequence 463, App
89	33.2	2.2	3226	9	US-10-124-819-463	Sequence 463, App
90	33.2	2.2	3226	9	US-10-124-822-463	Sequence 463, App
91	33.2	2.2	3226	9	US-10-140-925-463	Sequence 463, App
92	33.2	2.2	3226	9	US-10-160-498-463	Sequence 463, App

```

93 33.2 2.2 3226 9 US-09-989-328-352 Sequence 352, App
94 33.2 2.2 3226 9 US-09-993-583-352 Sequence 352, App
95 33.2 2.2 3226 9 US-10-121-041-463 Sequence 463, App
96 33.2 2.2 3226 9 US-10-121-043-463 Sequence 463, App
97 33.2 2.2 3226 9 US-10-121-043-463 Sequence 463, App
98 33.2 2.2 3226 9 US-10-123-215-463 Sequence 463, App
99 33.2 2.2 3226 9 US-10-123-902-463 Sequence 463, App
100 33.2 2.2 3226 9 US-10-123-908-463 Sequence 463, App
101 33.2 2.2 3226 9 US-10-123-909-463 Sequence 463, App
102 33.2 2.2 3226 9 US-10-123-910-463 Sequence 463, App
103 33.2 2.2 3226 9 US-10-124-813-463 Sequence 463, App
104 33.2 2.2 3226 9 US-10-124-817-463 Sequence 463, App
105 33.2 2.2 3226 9 US-10-124-824-463 Sequence 463, App
106 33.2 2.2 3226 9 US-10-125-922-463 Sequence 463, App
107 33.2 2.2 3226 9 US-10-125-924-463 Sequence 463, App
108 33.2 2.2 3226 9 US-10-127-825A-463 Sequence 463, App
109 33.2 2.2 3226 9 US-10-127-829A-463 Sequence 463, App
110 33.2 2.2 3226 9 US-10-127-835A-463 Sequence 463, App
111 33.2 2.2 3226 9 US-10-127-835A-463 Sequence 463, App
112 33.2 2.2 3226 9 US-10-127-835A-463 Sequence 463, App
113 33.2 2.2 3226 9 US-10-127-901A-463 Sequence 463, App
114 33.2 2.2 3226 9 US-10-128-693A-463 Sequence 463, App
115 33.2 2.2 3226 9 US-10-131-813A-463 Sequence 463, App
116 33.2 2.2 3226 9 US-10-131-818A-463 Sequence 463, App
117 33.2 2.2 3226 9 US-10-131-823A-463 Sequence 463, App
118 33.2 2.2 3226 9 US-10-131-824A-463 Sequence 463, App
119 33.2 2.2 3226 9 US-10-131-830A-463 Sequence 463, App
120 33.2 2.2 3226 9 US-10-131-837A-463 Sequence 463, App
121 33.2 2.2 3226 9 US-10-137-872A-463 Sequence 463, App
122 33.2 2.2 3226 9 US-10-140-860-463 Sequence 463, App
123 33.2 2.2 3226 9 US-10-142-417-463 Sequence 463, App
124 33.2 2.2 3226 9 US-10-147-500-463 Sequence 463, App
125 33.2 2.2 3226 9 US-10-147-502-463 Sequence 463, App
126 33.2 2.2 3226 9 US-10-147-515-463 Sequence 463, App
127 33.2 2.2 3226 9 US-10-147-517-463 Sequence 463, App
128 33.2 2.2 3226 9 US-10-147-519-463 Sequence 463, App
129 33.2 2.2 3226 9 US-10-147-526-463 Sequence 463, App
130 33.2 2.2 3226 9 US-10-147-527-463 Sequence 463, App
131 33.2 2.2 3226 9 US-10-152-395-463 Sequence 463, App
132 33.2 2.2 3226 9 US-10-157-782-463 Sequence 463, App
133 33.2 2.2 3226 9 US-09-941-922-352 Sequence 463, App
134 33.2 2.2 3226 9 US-09-992-521-352 Sequence 463, App
135 33.2 2.2 3226 9 US-10-121-040-463 Sequence 463, App
136 33.2 2.2 3226 9 US-10-121-056-463 Sequence 463, App
137 33.2 2.2 3226 9 US-10-121-051-463 Sequence 463, App
138 33.2 2.2 3226 9 US-10-123-235-463 Sequence 463, App
139 33.2 2.2 3226 9 US-10-124-818-463 Sequence 463, App
140 33.2 2.2 3226 9 US-10-125-926A-463 Sequence 463, App
141 33.2 2.2 3226 9 US-10-125-930A-463 Sequence 463, App
142 33.2 2.2 3226 9 US-10-127-831A-463 Sequence 463, App
143 33.2 2.2 3226 9 US-10-127-837A-463 Sequence 463, App
144 33.2 2.2 3226 9 US-10-127-838B-463 Sequence 463, App
145 33.2 2.2 3226 9 US-10-127-842A-463 Sequence 463, App
146 33.2 2.2 3226 9 US-10-127-843A-463 Sequence 463, App
147 33.2 2.2 3226 9 US-10-127-845A-463 Sequence 463, App
148 33.2 2.2 3226 9 US-10-127-846A-463 Sequence 463, App
149 33.2 2.2 3226 9 US-10-127-848A-463 Sequence 463, App
150 33.2 2.2 3226 9 US-10-127-849A-463 Sequence 463, App
151 33.2 2.2 3226 9 US-10-127-850A-463 Sequence 463, App

```

ALIGNMENTS

```

RESULT 1
US-09-884-889-7
; Sequence 7, Application us/09884889
; Patent No. US20020102680A1
; GENERAL INFORMATION:
; APPLICANT: DIVERSA CORPORATION
; APPLICANT: ROBERTSON, Dan
; APPLICANT: SANJAL, Indrajit
; APPLICANT: ADHIKARI, Robert
; TITLE OF INVENTION: CATALASES

```

```

; FILE REFERENCE: DIVER1100-4
; CURRENT APPLICATION NUMBER: US/09/884,889
; CURRENT FILING DATE: 2001-06-19
; PRIOR APPLICATION NUMBER: US 09/412,347
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: US 08/951,844
; PRIOR FILING DATE: 1997-10-16
; PRIOR APPLICATION NUMBER: US 08/674,887
; PRIOR FILING DATE: 1996-07-03
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 7
; LENGTH: 2238
; TYPE: DNA
; ORGANISM: Microscilla furvescens
; US-09-884-889-7

Query Match      24.0%   Score 357.6;   DB 10;   Length 2238;
Best Local Similarity 66.7%;   Pred. No. 6.4e-110;
Matches 542;   Conservative 0;   Mismatches 264;   Indels 6;   Gaps 2;

QY      683 ATCAATPCCCTGGGGGCTGATTTGATTATGCCACAGATTTCAACAGCTGGATATGA 742
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      168 ATCGGACCCAAACGACCCGATTTTGACTATGCCGAAGAGTTTAAAGCTAGATCTGC 227
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      743 GGCCTCAAAAAGATATCAACAGATTTGCTGCAACTCCAGGATTTGGTCCCTCGGA 802
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      228 AGCGTTTAAAGAGCCTGGCAGCGTAATGACAGATTCACAGGACTGGTGGCCAGACA 287
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      803 TTATGTCATATTATGGTCTTTTATGCTATGGCTTGGCAGGTCGCGGACGACATACAG 862
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      288 TTACGGTCAATTATGGCCCTTCTTTATACGATGGCTGGCAGCGCCGACCTTACC 347
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      863 GACATATGATGGCGGGAGGCGCCAGTGGTGTGACCAACGTTTGAACCGCTGAACAG 922
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      348 TATCGGTGATGGCCGCTGGTGGCGGTGCTCCGCTCACAGCGCTCGCGCTCTCAATAG 407
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      923 CTGGCCGGAATAAGCTTAATCTGGATAAAGCCGCTGATTTGCTGGCAGTCAAGAAAA 982
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      408 CTGGCCAGACAATGCCAATCTGGATAAAGCAGCTGCTCTTTGGCCCATCAACAAAA 467
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      983 ATACGCTCCAGTATTTCTGGGGAGACCTGATGTCCTGACTGTTGTAATTTGCCCTTGA 1042
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      468 ATACGCTCGAAAAATCTCTGGCGGATCTAATGATCTACAGGAAACGCTAGCTCTGA 527
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      1043 ATCCATGGGATTTAAACGCTGGGATTTGCTGGCGAAGAGAGATGCTGGAGTC--- 1099
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      528 AACTATGGCTTTAAACCTTTTGGTTTTCAGGTGGCAGAGCAGATGTATGGAGCCTGA 587
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      1100 GGACCTGTTATCTGGGGCTTGACAAACAGCCTTTCAGATAACCGGGATAA---AA 1156
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      588 AGAAGATGTATCTGGGAGCAGAAACCGAATGGCTGGGAGACAGCCTATGAGGTGA 647
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      1157 CGGGAACCTTCAGAACCTTCTTGGCGCCACGAGATGGGACTTATTTATGTCATCTCTGA 1216
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      648 CGGAGAGCTCGAAAAATCCCTGGGAGCCGTACAAATGGGACTCATCTATGTAACCCCGA 707
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      1217 AGGCCCCGGTGGAAACCAAGATCTCTGCTTCCGCGAAGATATCAGGAGAGCTTTTC 1276
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      708 AGGACCAACGCGCAGCCAGCCCTATCGCTGCTGCGGTGATATTCGTGAGACTTTGG 767
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      1277 ACGTATGCCATGATGATGAGGAGACCTGTCGCCCTGATCGCGGAGGAGGATATCTGG 1336
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      768 CGAATGGCAATGATGACGAAGAACCGTGGCTCTCATAGCGGTGGACACACCTTCGG 827
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      1337 TAAAGCACATGTCAGCGCTCTCTGAAAAATGTTTGGCGCAGGCTGATGTTGTCACC 1396
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      828 AAAAACCCATGCTGCTCCGATGCGGAGAAATGTTGGCGGAGAGCTGCGCGCAGG 887
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      1397 TGTGAGGAGCAGGAGCTGGGATGGAATAAATGTTGTTACAGAAACGCAATATAC 1456
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      888 TATTGAAGAAATGAGCCTGGGTGGAAACACACCTACGCGCACCGGTCGGGATAC 947
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

QY 1457 CATCACCAGTGGCTGGAAGGACCTGGTCTGA 1488
|||||
Db 948 CATCACCAGTGGACTAGAAAGGCGCTGGACCA 979

RESULT 2

US-09-884-889-5
; Sequence 5, Application US/09884889
; Patent No. US20020102680A1
; GENERAL INFORMATION:
; APPLICANT: DIVERSA CORPORATION
; APPLICANT: ROBERTSON, Dan
; APPLICANT: SANYAL, Indrajit
; APPLICANT: ADHIKARI, Robert
; APPLICANT: ADHIKARI, Robert
; TITLE OF INVENTION: CATALASES
; FILE REFERENCE: DIVER1100-4
; CURRENT APPLICATION NUMBER: US/09/884,889
; CURRENT FILING DATE: 2001-06-19
; PRIOR APPLICATION NUMBER: US 09/412,347
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: US 08/951,844
; PRIOR FILING DATE: 1997-10-16
; PRIOR APPLICATION NUMBER: US 08/674,887
; PRIOR FILING DATE: 1996-07-03
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5
; LENGTH: 2262
; TYPE: DNA
; ORGANISM: Alcaligenes (Deleya) aquamarinus
US-09-884-889-5

Query Match 15.1%; Score 224.6; DB 10; Length 2262;
Best Local Similarity 57.4%; Pred. No. 9.2e-65;
Matches 476; Conservative 0; Mismatches 329; Indels 24; Gaps 3;

QY 682 AATCAATCCCTGGGGGCTGATTTGATATGCCACAGATTTCAACAGCTGGATGG 741
|||||
Db 218 AATCAGACCGATGGATTCGGATTTCAACTACCGTGAAGAAGTACCAAGCTCGNTTCG 277
|||||
QY 742 AGCTCTGAAAAGATATCAAGATTTGCTGACAACTTCCCAAGGATTTGGCTGCGG 801
|||||
Db 278 ACCGCTGAAGAAGATGTCACGCGTGTATGACCGATACCAAGAGTGGTGGCCGCTG 337
|||||
QY 802 ATTATGGTCATTATGTCCTTTCTTATTCGTATGCTGGACGCTGGCGGAACATACA 861
|||||
Db 338 ACTGGGGCACTACGCGGTTTGTATCGTATCGTGGCTGGCACTCGCTGGCACCTACC 397
|||||
QY 862 GGACATATGATGGCGGGGAGCGCCAGTGGTGGTCAACACGTTTGAACCGCTGAACA 921
|||||
Db 398 GTATTGCTGATGGCGGCGGGGCGGTGTACCGGAAGCCAGCGCTTGGCACCGCTCAACT 457
|||||
QY 922 GTGGCGGATACGTTAATCTGGATAAAGCCCGCTGATTTGCTGGCCAGTCAAGAAA 981
|||||
Db 458 CTGGCGGACACGTCAGCTGATTAAGCGCGCTGCTGTGGCCGATCAAGAAGA 517
|||||
QY 982 AATACGGCTCCAGTATTTCTGGGAGACCTGATGTCCTGATGTAATGTTGCCCTTG 1041
|||||
Db 518 AGTACGGCAACAAATCAGCTGGCGAGACCTGATGATTTCTGGCTGGCAGCGCTGATG 577
|||||
QY 1042 AATCCATGGATTTAAACCGCTGGATTTCTGGCGGAAGAGATGACTGGAGTCGG 1101
|||||
Db 578 AGTCCATGGCTTACCTGCTTACGGCTTCTTTTCGGCGCGCTGATATTTGGGAACCGG 637
|||||
QY 1102 ACCTGG---TATACCTGGGCTTGACAACAGCTCTTGAGAT----- 1142
|||||
Db 638 AAAAGATATCTACTGGGTGACGAAGAAGAGTGGCTGGCACCTTCTGACCAACGCTACG 697
|||||
QY 1143 --AACCGGATAAAGCGGAACCTTCAAGAACTTTCAGAACTTTCGCGCCACGAGATGGACTTA 1200
|||||
Db 698 GCGACGTGAACAGCCAGACACCATGGAACACCGCTGGCGGCTGTCCTCAATGGGCTGA 757
|||||
QY 1201 TTTATGTCATCTCTGAAGCGCGGTGGAAACCAACATGCTCTGCGCTTCGCGCAAGATA 1260
|||||

Db 758 TCTATGTGAACCGGAAGGTGTTAAACGGCCACCTGATCGCTGAGAACCGCACACGAG 817
|||||
QY 1261 TCAGGGAAGCTTTTTCACGCTATGGCATGATGAGGAGACTGTGGCCCTCATCGCGG 1320
|||||
Db 818 TACTTGAACCTTCGCGCGGTATGCGCATGAACGCAAGAAACCGCGCCTCACAGCTG 877
|||||
QY 1321 GAGGCGATACATTTGTTAAAGCACATGGTGACGCTCTCTGAAAAATGTTTGGCGCAG 1380
|||||
Db 878 GCGGCCACACCGTCCGCTAATTTGTACGTAATGCAATGCTCTGCGT---TAGCCCTG 934
|||||
QY 1381 GGCCTGATGTCACCTGTGGAGGACGAGCTGGATGGAGAAATAAATGTTGGTACAG 1440
|||||
Db 935 ACCCAAAAGCCTCTGACGTTGAAAAACGAGGCTTAGGTTGGGCAACCCCAACATCGAG 994
|||||
QY 1441 GAAACGCGCAAAATATACCATCACCAGTGGCTTGAAGAGAGCTGCTGCTGAC 1489
|||||
Db 995 GCAAGGCAACGACCGCTGACCTCGGTATCAAGGTGCTTGGACCAC 1043
|||||

RESULT 3

US-10-123-155-254
; Sequence 254, Application US/10123155
; Publication No. US20030068794A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330RIC30
; CURRENT APPLICATION NUMBER: US/10/123,155
; CURRENT FILING DATE: 2002-04-15
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 254
; LENGTH: 567
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-123-155-254

Query Match 2.6%; Score 38.6; DB 9; Length 567;
Best Local Similarity 9.2%; Pred. No. 0.047;
Matches 41; Conservative 134; Mismatches 270; Indels 0; Gaps 0;
QY 830 TCGTATGGCTGGCAGCGTCCCGAACATACAGACATATATGATGGCGGGGAGCGCCAG 889
|||||
Db 52 TEGADDDGKGPSIMDFVTHSGKVLGNEDTADVDGYKQVEDIILLRELHVNHYRFS 111
|||||
QY 890 TGTGTGTGACAGCTTTTGAACCGCTGAACAGCTGGCCGAGTAACGTTAATCTGGATA 949
|||||
Db 112 LSWPRLLPGIRAEQVKNKIEFYSLDILLSSNTPITVLHWDLPQLLVKYGWQV 171
|||||
QY 950 AGCCCGTCGATGCTGTGGCCAGTCAAGAAAAATACGGCTCCAGTATTTCTCTGGGAGA 1009
|||||
Db 172 VSMANFRDYNALCFEAFGDRVKHWITFSDPRMAEKGYETGHHAPGLKLRGTGLYKAH 231
|||||
QY 1010 CCTGATGGCTCCTGACCTGGTGAATGTTGCCCTTGAATCCATGGGATTTAAACCGTGGGATT 1069
|||||

Db 426 ILEKEGVKILTPPEGEIFREPYLDIHKLVSQGIKGGDERGLLSLAFHPNKKNGKLYVSY 485
Qy 439 CTGTAGTATCTATCCAGCAATAGTATATCTCTGTTGCATCAATAAGTTGACTTTTGA 498
Db 486 TTQERWAGPHDIIRVVEYTVSRKNPHOVDLRTARVLEVAELHRKHLGQLFGPDG 545
Qy 499 TACAACATGGAATTTCCCTTAATCGGAGCTATTGATGATAAATAAACTCTTCTG 558
Db 546 FLYIILGDMITLDDMEEDGLSDFGVLRLDVTDMCNVYSIPRPNHNSNQPP 605
Qy 559 TCTGATCTCTGCGGCTATCGGGAGC 587
Db 606 VFAHGLHDPGRCAVDRHPTDININLTLC 634

RESULT 19

US-10-092-154-1415/c
; Sequence 1415, Application US/10092154
; Publication No. US20030054375A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC009C1
; CURRENT APPLICATION NUMBER: US/10/092,154
; CURRENT FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 2003
; Prior Application removed - See File Wrapper or Palm
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1415
; LENGTH: 6383
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-092-154-1415

Query Match 2.2%; Score 33.4; DB 9; Length 6383;
Best Local Similarity 55.7%; Pred. No. 16;
Matches 64; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

Qy 1303 CTGTGGCCCTGATCGGGAGGCGCATATTGTTAAAGCACATGGTGCGAGCTCTCTG 1362
Db 325 CTGCAGGTGTATCGCGGTGAGGCCAGTGAGATGCACACGCGGGCAGCTGCTCTG 266
Qy 1363 AAAAATGATTTGGCGAGGCGCTGATGTGTGACCTGTGGAGGACGAGGACTGGG 1417
Db 265 CAGAACGGCTTTGCTGACATCATCCCGTGACCAGCAGAAGGTGGAGGATGTGG 211

RESULT 20

US-09-764-847-1415/c
; Sequence 1415, Application US/09764847
; Patent No. US20020132767A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC009
; CURRENT APPLICATION NUMBER: US/09/764,847
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 2003
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1415
; LENGTH: 6383
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-847-1415

Query Match 2.2%; Score 33.4; DB 10; Length 6383;
Best Local Similarity 55.7%; Pred. No. 16;
Matches 64; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

Qy 1303 CTGTGGCCCTGATCGGGAGGCGCATATTGTTAAAGCACATGGTGCGAGCTCTCTG 1362
Db 325 CTGCAGGTGTATCGCGGTGAGGCCAGTGAGATGCACACGCGGGCAGCTGCTCTG 266
Qy 1363 AAAAATGATTTGGCGAGGCGCTGATGTGTGACCTGTGGAGGACGAGGACTGGG 1417
Db 265 CAGAACGGCTTTGCTGACATCATCCCGTGACCAGCAGAAGGTGGAGGATGTGG 211

Db 325 CTGCAGGTGTATCGGGAGGCGCGAGTGAGATGCACACGCGGCGAGCTGCTCTG 266
Qy 1363 AAAAATGATTTGGCGAGGCGCTGATGTGTGACCTGTGGAGGACGAGGACTGGG 1417
Db 265 CAGAACGGCTTTGCTGACATCATCCCGTGACCAGCAGAAGGTGGAGGATGTGG 211

RESULT 21

US-10-092-154-1414/c
; Sequence 1414, Application US/10092154
; Publication No. US20030054375A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC009C1
; CURRENT APPLICATION NUMBER: US/10/092,154
; CURRENT FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 2003
; Prior Application removed - See File Wrapper or Palm
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1414
; LENGTH: 21732
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-092-154-1414

Query Match 2.2%; Score 33.4; DB 9; Length 21732;
Best Local Similarity 55.7%; Pred. No. 37;
Matches 64; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

Qy 1303 CTGTGGCCCTGATCGGGAGGCGCATATTGTTAAAGCACATGGTGCGAGCTCTCTG 1362
Db 15674 CTGCAGGTGTATCGCGGTGAGGCCAGTGAGATGCACACGCGGGCAGCTGCTCTG 15615
Qy 1363 AAAAATGATTTGGCGAGGCGCTGATGTGTGACCTGTGGAGGACGAGGACTGGG 1417
Db 15614 CAGAACGGCTTTGCTGACATCATCCCGTGACCAGCAGAAGGTGGAGGATGTGG 15560

RESULT 22

US-09-764-847-1414/c
; Sequence 1414, Application US/09764847
; Patent No. US20020132767A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC009
; CURRENT APPLICATION NUMBER: US/09/764,847
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 2003
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1414
; LENGTH: 21732
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-847-1414

Query Match 2.2%; Score 33.4; DB 10; Length 21732;
Best Local Similarity 55.7%; Pred. No. 37;
Matches 64; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

Qy 1303 CTGTGGCCCTGATCGGGAGGCGCATATTGTTAAAGCACATGGTGCGAGCTCTCTG 1362
Db 15674 CTGCAGGTGTATCGCGGTGAGGCCAGTGAGATGCACACGCGGGCAGCTGCTCTG 15615
Qy 1363 AAAAATGATTTGGCGAGGCGCTGATGTGTGACCTGTGGAGGACGAGGACTGGG 1417
Db 15614 CAGAACGGCTTTGCTGACATCATCCCGTGACCAGCAGAAGGTGGAGGATGTGG 15560

RESULT 23

US-09-918-995-29551

```
; Sequence 29551, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 29551
; LENGTH: 484
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(484)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-29551

Query Match          2.28; Score 33.2; DB 9; Length 484;
Best Local Similarity 57.88; Pred. No. 2.9;
Matches 59; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

QY 1316 CGCGGGAGGCATACATTTGGTAAAGCAGCATGTGTCAGCGCTCTCTCTGAAAAATGTAATGG 1375
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 149 CGGTGGAGGCCAGTGAGATGGCAGCACACGGGGCAGCTGGTCTCTCAGACAGCGCTTG 208
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1376 CGCAGGCCCTGTGGTGCACCTGTGGAGGAGCAGGAGCTGGG 1417
    | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 209 CTGACATCATCACCGTGTACCAGCAGAAAGTGGAGGATGTGG 250

RESULT 24
US-09-992-598-352
; Sequence 352, Application US/09992598
; Patent No. US20020160384A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2730PIC20
; CURRENT APPLICATION NUMBER: US/09/992,598
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
```

,	PRIOR APPLICATION NUMBER:	60/089512
,	PRIOR FILING DATE:	1998-06-16
,	PRIOR APPLICATION NUMBER:	60/089514
,	PRIOR FILING DATE:	1998-06-16
,	PRIOR APPLICATION NUMBER:	60/089532
,	PRIOR FILING DATE:	1998-06-17
,	PRIOR APPLICATION NUMBER:	60/089538
,	PRIOR FILING DATE:	1998-06-17
,	PRIOR APPLICATION NUMBER:	60/089598
,	PRIOR FILING DATE:	1998-06-17
,	PRIOR APPLICATION NUMBER:	60/089599
,	PRIOR FILING DATE:	1998-06-17
,	PRIOR APPLICATION NUMBER:	60/089600
,	PRIOR FILING DATE:	1998-06-17
,	PRIOR APPLICATION NUMBER:	60/089653
,	PRIOR FILING DATE:	1998-06-17
,	PRIOR APPLICATION NUMBER:	60/089801
,	PRIOR FILING DATE:	1998-06-18
,	PRIOR APPLICATION NUMBER:	60/089907
,	PRIOR FILING DATE:	1998-06-18
,	PRIOR APPLICATION NUMBER:	60/089908
,	PRIOR FILING DATE:	1998-06-18
,	PRIOR APPLICATION NUMBER:	60/089947
,	PRIOR FILING DATE:	1998-06-19
,	PRIOR APPLICATION NUMBER:	60/089948
,	PRIOR FILING DATE:	1998-06-19
,	PRIOR APPLICATION NUMBER:	60/089952
,	PRIOR FILING DATE:	1998-06-19
,	PRIOR APPLICATION NUMBER:	60/090246
,	PRIOR FILING DATE:	1998-06-22
,	PRIOR APPLICATION NUMBER:	60/090252
,	PRIOR FILING DATE:	1998-06-22
,	PRIOR APPLICATION NUMBER:	60/090254
,	PRIOR FILING DATE:	1998-06-22
,	PRIOR APPLICATION NUMBER:	60/090349
,	PRIOR FILING DATE:	1998-06-23
,	PRIOR APPLICATION NUMBER:	60/090355
,	PRIOR FILING DATE:	1998-06-23
,	PRIOR APPLICATION NUMBER:	60/090429
,	PRIOR FILING DATE:	1998-06-24
,	PRIOR APPLICATION NUMBER:	60/090431
,	PRIOR FILING DATE:	1998-06-24
,	PRIOR APPLICATION NUMBER:	60/090435
,	PRIOR FILING DATE:	1998-06-24
,	PRIOR APPLICATION NUMBER:	60/090444
,	PRIOR FILING DATE:	1998-06-24
,	PRIOR APPLICATION NUMBER:	60/090445
,	PRIOR FILING DATE:	1998-06-24
,	PRIOR APPLICATION NUMBER:	60/090472
,	PRIOR FILING DATE:	1998-06-24
,	PRIOR APPLICATION NUMBER:	60/090535
,	PRIOR FILING DATE:	1998-06-24
,	PRIOR APPLICATION NUMBER:	60/090540
,	PRIOR FILING DATE:	1998-06-24
,	PRIOR APPLICATION NUMBER:	60/090542
,	PRIOR FILING DATE:	1998-06-24
,	PRIOR APPLICATION NUMBER:	60/090557
,	PRIOR FILING DATE:	1998-06-24
,	PRIOR APPLICATION NUMBER:	60/090676
,	PRIOR FILING DATE:	1998-06-25
,	PRIOR APPLICATION NUMBER:	60/090678
,	PRIOR FILING DATE:	1998-06-25
,	PRIOR APPLICATION NUMBER:	60/090690
,	PRIOR FILING DATE:	1998-06-25
,	PRIOR APPLICATION NUMBER:	60/090694
,	PRIOR FILING DATE:	1998-06-25
,	PRIOR APPLICATION NUMBER:	60/090695
,	PRIOR FILING DATE:	1998-06-25
,	PRIOR APPLICATION NUMBER:	60/090696
,	PRIOR FILING DATE:	1998-06-25
,	PRIOR APPLICATION NUMBER:	60/090862
,	PRIOR FILING DATE:	1998-06-26
,	PRIOR APPLICATION NUMBER:	60/090863

?	PRIOR FILING DATE:	1998-06-26
?	PRIOR APPLICATION NUMBER:	60/091360
?	PRIOR FILING DATE:	1998-07-01
?	PRIOR APPLICATION NUMBER:	60/091478
?	PRIOR FILING DATE:	1998-07-02
?	PRIOR APPLICATION NUMBER:	60/091544
?	PRIOR FILING DATE:	1998-07-01
?	PRIOR APPLICATION NUMBER:	60/091519
?	PRIOR FILING DATE:	1998-07-02
?	PRIOR APPLICATION NUMBER:	60/091626
?	PRIOR FILING DATE:	1998-07-02
?	PRIOR APPLICATION NUMBER:	60/091633
?	PRIOR FILING DATE:	1998-07-02
?	PRIOR APPLICATION NUMBER:	60/091978
?	PRIOR FILING DATE:	1998-07-07
?	PRIOR APPLICATION NUMBER:	60/091982
?	PRIOR FILING DATE:	1998-07-07
?	PRIOR APPLICATION NUMBER:	60/092182
?	PRIOR FILING DATE:	1998-07-09

	Query Match	2.2%	Score 33.2;	DB 9;	Length 3226;
	Best Local Similarity	54.0%;	Pred. No. 11;		
	Matches 68;	Conservative 0;	Mismatches 58;	Indels 0;	Gaps 0;
QY	1355	GTCTCTGAAAAATGATTCGGCCGAGGCCCTGTATGGTGCACTGTGTGGAGGACGAGGACT	1414		
Db	1998	GTCTCATTTACAATGCATTTTCAGTCTGTGAGCATTTGGGAAGCTGTGCATTTGAAAAGGCGCT	2057		
QY	1415	GGGATGGAAAAATTAATTTGGTGTACAGAAACGGCAATATACCATCACCAGTGGCCTGGA	1474		
Db	2058	TGGATTTATCCCTGTACTTTGAACATGAAACTGAAATTTATGGCCGCTGTTTCAAGGTTTGA	2117		
QY	1475	AGGAGC	1480		
Db	2118	ATGAGC	2123		

RESULT 25
US-09-989-293A-352
Sequence 352, Application US/09989293A
Patent No. US2002017716A1
GENERAL INFORMATION:
APPLICANT: Ashkenazi, Avi J.
APPLICANT: Baker, Kevin P.
APPLICANT: Botstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Eaton, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Fong, Sherman
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Kljavin, Ivar J.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K.
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: Secured and Trans-
Acids Encoding the
FILE REFERENCE: P2730PIC66
CURRENT APPLICATION NUMBER: US/09/989;
CURRENT FILING DATE: 2001-11-20
PRIOR APPLICATION NUMBER: 60/049787
PRIOR FILING DATE: 1997-06-16

1	PRIOR FILING DATE: 1998-06-16	
2	PRIOR APPLICATION NUMBER: 60/089512	
3	PRIOR FILING DATE: 1998-06-16	
4	PRIOR APPLICATION NUMBER: 60/089514	
5	PRIOR FILING DATE: 1998-06-16	
6	PRIOR APPLICATION NUMBER: 60/089532	
7	PRIOR FILING DATE: 1998-06-17	
8	PRIOR APPLICATION NUMBER: 60/089538	
9	PRIOR FILING DATE: 1998-06-17	
10	PRIOR APPLICATION NUMBER: 60/089598	
11	PRIOR FILING DATE: 1998-06-17	
12	PRIOR APPLICATION NUMBER: 60/089599	
13	PRIOR FILING DATE: 1998-06-17	
14	PRIOR APPLICATION NUMBER: 60/089600	
15	PRIOR FILING DATE: 1998-06-17	
16	PRIOR APPLICATION NUMBER: 60/089653	
17	PRIOR FILING DATE: 1998-06-17	
18	PRIOR APPLICATION NUMBER: 60/089801	
19	PRIOR FILING DATE: 1998-06-18	
20	PRIOR APPLICATION NUMBER: 60/089907	
21	PRIOR FILING DATE: 1998-06-18	
22	PRIOR APPLICATION NUMBER: 60/089908	
23	PRIOR FILING DATE: 1998-06-18	
24	PRIOR APPLICATION NUMBER: 60/089947	
25	PRIOR FILING DATE: 1998-06-19	
26	PRIOR APPLICATION NUMBER: 60/089948	
27	PRIOR FILING DATE: 1998-06-19	
28	PRIOR APPLICATION NUMBER: 60/089952	
29	PRIOR FILING DATE: 1998-06-19	
30	PRIOR APPLICATION NUMBER: 60/090246	
31	PRIOR FILING DATE: 1998-06-22	
32	PRIOR APPLICATION NUMBER: 60/090252	
33	PRIOR FILING DATE: 1998-06-22	
34	PRIOR APPLICATION NUMBER: 60/090254	
35	PRIOR FILING DATE: 1998-06-22	
36	PRIOR APPLICATION NUMBER: 60/090349	
37	PRIOR FILING DATE: 1998-06-23	
38	PRIOR APPLICATION NUMBER: 60/090355	
39	PRIOR FILING DATE: 1998-06-23	
40	PRIOR APPLICATION NUMBER: 60/090429	
41	PRIOR FILING DATE: 1998-06-24	
42	PRIOR APPLICATION NUMBER: 60/090444	
43	PRIOR FILING DATE: 1998-06-24	
44	PRIOR APPLICATION NUMBER: 60/090445	
45	PRIOR FILING DATE: 1998-06-24	
46	PRIOR APPLICATION NUMBER: 60/090472	
47	PRIOR FILING DATE: 1998-06-24	
48	PRIOR APPLICATION NUMBER: 60/090535	
49	PRIOR FILING DATE: 1998-06-24	
50	PRIOR APPLICATION NUMBER: 60/090540	
51	PRIOR FILING DATE: 1998-06-24	
52	PRIOR APPLICATION NUMBER: 60/090542	
53	PRIOR FILING DATE: 1998-06-24	
54	PRIOR APPLICATION NUMBER: 60/090557	
55	PRIOR FILING DATE: 1998-06-24	
56	PRIOR APPLICATION NUMBER: 60/090676	
57	PRIOR FILING DATE: 1998-06-25	
58	PRIOR APPLICATION NUMBER: 60/090678	
59	PRIOR FILING DATE: 1998-06-25	
60	PRIOR APPLICATION NUMBER: 60/090690	
61	PRIOR FILING DATE: 1998-06-25	
62	PRIOR APPLICATION NUMBER: 60/090694	
63	PRIOR FILING DATE: 1998-06-25	
64	PRIOR APPLICATION NUMBER: 60/090695	
65	PRIOR FILING DATE: 1998-06-25	
66	PRIOR APPLICATION NUMBER: 60/090696	
67	PRIOR FILING DATE: 1998-06-25	
68	PRIOR APPLICATION NUMBER: 60/090862	
69	PRIOR FILING DATE: 1998-06-26	

```

; PRIOR APPLICATION NUMBER: 60/090863
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/091360
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091478
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091544
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091626
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091633
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091978
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/092182
; PRIOR FILING DATE: 1998-07-09

Query Match      2.2%; Score 33.2; DB 9; Length 3226;
Best Local Similarity 54.0%; Pred. No. 11;
Matches 68; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY 1355 GTCCTCGTAAATGTTGGCGCAGGCGCTGATGTCGACCTGTGGAGGACGAGGACT 1414
      ||||| | ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1998 GTCCTATTACATGCTTCAGCTCGTCGAGCATTTGGGAAGCTGTCCATTGAAAAGGCT 2057
      ||||| | ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1415 GGGATGGAATAATATGTGTGTACAGGAAAGCGCAATATACCATCACCAGTGGCCTGGA 1474
      ||||| | ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2058 TGGATTATCCCTGTACTTGAACATGAACATGAATATATGCCCGTGTTCAGGTTTGA 2117
      ||||| | ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1475 AGGAGC 1480
      |||||
Db 2118 ATGAGC 2123

RESULT 26
US-09-989-735-352
; Sequence 352, Application US/09989735
; Publication No. US20020193299A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tamas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2730P1C61
; CURRENT FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/075945
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/087106
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087609
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087827
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: 60/088021
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088025
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088026
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088028
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088029
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088030
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088033
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088326
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088167
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088202
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088212
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088217
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088655
; PRIOR FILING DATE: 1998-06-09
; PRIOR APPLICATION NUMBER: 60/088734
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088738
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088742
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088810
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088824
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088826
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088858
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088861
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088876
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/089105
; PRIOR FILING DATE: 1998-06-12
```


1	PRIOR APPLICATION NUMBER: 60/089100
2	PRIOR FILING DATE: 1998-06-12
3	PRIOR APPLICATION NUMBER: 60/089440
4	PRIOR FILING DATE: 1998-06-16
5	PRIOR APPLICATION NUMBER: 60/089512
6	PRIOR FILING DATE: 1998-06-16
7	PRIOR APPLICATION NUMBER: 60/089514
8	PRIOR FILING DATE: 1998-06-16
9	PRIOR APPLICATION NUMBER: 60/089532
10	PRIOR FILING DATE: 1998-06-17
11	PRIOR APPLICATION NUMBER: 60/089538
12	PRIOR FILING DATE: 1998-06-17
13	PRIOR APPLICATION NUMBER: 60/089598
14	PRIOR FILING DATE: 1998-06-17
15	PRIOR APPLICATION NUMBER: 60/089599
16	PRIOR FILING DATE: 1998-06-17
17	PRIOR APPLICATION NUMBER: 60/089600
18	PRIOR FILING DATE: 1998-06-17
19	PRIOR APPLICATION NUMBER: 60/089655
20	PRIOR FILING DATE: 1998-06-17
21	PRIOR APPLICATION NUMBER: 60/089801
22	PRIOR FILING DATE: 1998-06-18
23	PRIOR APPLICATION NUMBER: 60/089907
24	PRIOR FILING DATE: 1998-06-18
25	PRIOR APPLICATION NUMBER: 60/089908
26	PRIOR FILING DATE: 1998-06-18
27	PRIOR APPLICATION NUMBER: 60/089947
28	PRIOR FILING DATE: 1998-06-19
29	PRIOR APPLICATION NUMBER: 60/089948
30	PRIOR FILING DATE: 1998-06-19
31	PRIOR APPLICATION NUMBER: 60/089952
32	PRIOR FILING DATE: 1998-06-19
33	PRIOR APPLICATION NUMBER: 60/090246
34	PRIOR FILING DATE: 1998-06-22
35	PRIOR APPLICATION NUMBER: 60/090252
36	PRIOR FILING DATE: 1998-06-22
37	PRIOR APPLICATION NUMBER: 60/090254
38	PRIOR FILING DATE: 1998-06-22
39	PRIOR APPLICATION NUMBER: 60/090349
40	PRIOR FILING DATE: 1998-06-23
41	PRIOR APPLICATION NUMBER: 60/090355
42	PRIOR FILING DATE: 1998-06-23
43	PRIOR APPLICATION NUMBER: 60/090429
44	PRIOR FILING DATE: 1998-06-24
45	PRIOR APPLICATION NUMBER: 60/090431
46	PRIOR FILING DATE: 1998-06-24
47	PRIOR APPLICATION NUMBER: 60/090435
48	PRIOR FILING DATE: 1998-06-24
49	PRIOR APPLICATION NUMBER: 60/090444
50	PRIOR FILING DATE: 1998-06-24
51	PRIOR APPLICATION NUMBER: 60/090445
52	PRIOR FILING DATE: 1998-06-24
53	PRIOR APPLICATION NUMBER: 60/090472
54	PRIOR FILING DATE: 1998-06-24
55	PRIOR APPLICATION NUMBER: 60/090535
56	PRIOR FILING DATE: 1998-06-24
57	PRIOR APPLICATION NUMBER: 60/090540
58	PRIOR FILING DATE: 1998-06-24
59	PRIOR APPLICATION NUMBER: 60/090542
60	PRIOR FILING DATE: 1998-06-24
61	PRIOR APPLICATION NUMBER: 60/090557
62	PRIOR FILING DATE: 1998-06-24
63	PRIOR APPLICATION NUMBER: 60/090676
64	PRIOR FILING DATE: 1998-06-25
65	PRIOR APPLICATION NUMBER: 60/090678
66	PRIOR FILING DATE: 1998-06-25
67	PRIOR APPLICATION NUMBER: 60/090690
68	PRIOR FILING DATE: 1998-06-25
69	PRIOR APPLICATION NUMBER: 60/090694
70	PRIOR FILING DATE: 1998-06-25
71	PRIOR APPLICATION NUMBER: 60/090695
72	PRIOR FILING DATE: 1998-06-25
73	PRIOR APPLICATION NUMBER: 60/090696

[illegible]

RESULT 29

```

US-09-990-436-352
; Sequence 352, Application US/09990436
; Publication No. US20020198148A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnovers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Collin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Trans
; TITLE OF INVENTION: Acids Encoding 14
; FILE REFERENCE: P273081C14

```

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	8												

1	PRIOR FILING DATE: 1998-06-11
2	PRIOR APPLICATION NUMBER: 60/089105
3	PRIOR FILING DATE: 1998-06-12
4	PRIOR APPLICATION NUMBER: 60/089440
5	PRIOR FILING DATE: 1998-06-16
6	PRIOR APPLICATION NUMBER: 60/089512
7	PRIOR FILING DATE: 1998-06-16
8	PRIOR APPLICATION NUMBER: 60/089514
9	PRIOR FILING DATE: 1998-06-16
10	PRIOR APPLICATION NUMBER: 60/089532
11	PRIOR FILING DATE: 1998-06-17
12	PRIOR APPLICATION NUMBER: 60/089538
13	PRIOR FILING DATE: 1998-06-17
14	PRIOR APPLICATION NUMBER: 60/089598
15	PRIOR FILING DATE: 1998-06-17
16	PRIOR APPLICATION NUMBER: 60/089599
17	PRIOR FILING DATE: 1998-06-17
18	PRIOR APPLICATION NUMBER: 60/089600
19	PRIOR FILING DATE: 1998-06-17
20	PRIOR APPLICATION NUMBER: 60/089653
21	PRIOR FILING DATE: 1998-06-17
22	PRIOR APPLICATION NUMBER: 60/089801
23	PRIOR FILING DATE: 1998-06-18
24	PRIOR APPLICATION NUMBER: 60/089907
25	PRIOR FILING DATE: 1998-06-18
26	PRIOR APPLICATION NUMBER: 60/089908
27	PRIOR FILING DATE: 1998-06-18
28	PRIOR APPLICATION NUMBER: 60/089947
29	PRIOR FILING DATE: 1998-06-19
30	PRIOR APPLICATION NUMBER: 60/089948
31	PRIOR FILING DATE: 1998-06-19
32	PRIOR APPLICATION NUMBER: 60/089952
33	PRIOR FILING DATE: 1998-06-19
34	PRIOR APPLICATION NUMBER: 60/090246
35	PRIOR FILING DATE: 1998-06-22
36	PRIOR APPLICATION NUMBER: 60/090252
37	PRIOR FILING DATE: 1998-06-22
38	PRIOR APPLICATION NUMBER: 60/090254
39	PRIOR FILING DATE: 1998-06-22
40	PRIOR APPLICATION NUMBER: 60/090349
41	PRIOR FILING DATE: 1998-06-23
42	PRIOR APPLICATION NUMBER: 60/090355
43	PRIOR FILING DATE: 1998-06-23
44	PRIOR APPLICATION NUMBER: 60/090429
45	PRIOR FILING DATE: 1998-06-24
46	PRIOR APPLICATION NUMBER: 60/090431
47	PRIOR FILING DATE: 1998-06-24
48	PRIOR APPLICATION NUMBER: 60/090435
49	PRIOR FILING DATE: 1998-06-24
50	PRIOR APPLICATION NUMBER: 60/090444
51	PRIOR FILING DATE: 1998-06-24
52	PRIOR APPLICATION NUMBER: 60/090445
53	PRIOR FILING DATE: 1998-06-24
54	PRIOR APPLICATION NUMBER: 60/090472
55	PRIOR FILING DATE: 1998-06-24
56	PRIOR APPLICATION NUMBER: 60/090535
57	PRIOR FILING DATE: 1998-06-24
58	PRIOR APPLICATION NUMBER: 60/090540
59	PRIOR FILING DATE: 1998-06-24
60	PRIOR APPLICATION NUMBER: 60/090542
61	PRIOR FILING DATE: 1998-06-24
62	PRIOR APPLICATION NUMBER: 60/090557
63	PRIOR FILING DATE: 1998-06-24
64	PRIOR APPLICATION NUMBER: 60/090676
65	PRIOR FILING DATE: 1998-06-25
66	PRIOR APPLICATION NUMBER: 60/090678
67	PRIOR FILING DATE: 1998-06-25
68	PRIOR APPLICATION NUMBER: 60/090690
69	PRIOR FILING DATE: 1998-06-25
70	PRIOR APPLICATION NUMBER: 60/090694
71	PRIOR FILING DATE: 1998-06-25
72	PRIOR APPLICATION NUMBER: 60/090695
73	PRIOR FILING DATE: 1998-06-25

```

: PRIOR APPLICATION NUMBER: 60/090696
: PRIOR FILING DATE: 1998-06-25
: PRIOR APPLICATION NUMBER: 60/090862
: PRIOR FILING DATE: 1998-06-26
: PRIOR APPLICATION NUMBER: 60/090863
: PRIOR FILING DATE: 1998-06-26
: PRIOR APPLICATION NUMBER: 60/091360
: PRIOR FILING DATE: 1998-07-01
: PRIOR APPLICATION NUMBER: 60/091478
: PRIOR FILING DATE: 1998-07-02
: PRIOR APPLICATION NUMBER: 60/091544
: PRIOR FILING DATE: 1998-07-01
: PRIOR APPLICATION NUMBER: 60/091519
: PRIOR FILING DATE: 1998-07-02
: PRIOR APPLICATION NUMBER: 60/091626
: PRIOR FILING DATE: 1998-07-02
: PRIOR APPLICATION NUMBER: 60/091633
: PRIOR FILING DATE: 1998-07-02
: PRIOR APPLICATION NUMBER: 60/091978
: PRIOR FILING DATE: 1998-07-07
: PRIOR APPLICATION NUMBER: 60/091982
: PRIOR FILING DATE: 1998-07-07
: PRIOR APPLICATION NUMBER: 60/092182
: PRIOR FILING DATE: 1998-07-09

```

Query Match	2.2%	Score 33.2;	DB 9;	Length 3226;
Best Local Similarity	54.0%;	Pred. No. 11;		
Matches 68;	Conservative 0;	Mismatches 58;	Indels 0;	Gaps 0;
QY	1355	GTCCTCCGAAAAATGATTGGCCAGGGCCCTGATGGTGCACTGTGGAGGACGGGACT	1414	
Db	1998	GTCCTCATTAACATGTGCATTTTCAGCTGCTGAGCATTTGGGAAGCTGCCATTGAAAAAGSCCT	2057	
QY	1415	GGGATGGAAAAATAAATGTGTGTACAGGAAACGGCAAAATATACCATCACCCTGGCCCTGGGA	1474	
Db	2058	TGGATTATCCCTGTACTTGAACACATGAACACTGAAATATTGCCCGTGTTCACAGGTTTGA	2117	
QY	1475	AGGAGC	1480	
Db	2118	ATGAGC	2123	

RESULT 30

US-09-991-181-352

Sequence 352, Application US/09991181

Publication No. US20020197615A1

GENERAL INFORMATION:

APPLICANT: Ashkenazi, Avi J.

APPLICANT: Baker, Kevin P.

APPLICANT: Botstein, David

APPLICANT: Desnovers, Luc

APPLICANT: Eaton, Dan L.

APPLICANT: Ferrara, Napoleone

APPLICANT: Fong, Sherman

APPLICANT: Gerber, Hanspeter

APPLICANT: Gerritsen, Mary E.

APPLICANT: Goddard, Audrey

APPLICANT: Godowski, Paul J.

APPLICANT: Grimaldi, J. Christopher

APPLICANT: Gurney, Austin L.

APPLICANT: Kijavini, Ivar J.

APPLICANT: Napier, Mary A.

APPLICANT: Pan, James

APPLICANT: Paoni, Nicholas F.

APPLICANT: Roy, Margaret Ann

APPLICANT: Stewart, Timothy A.

APPLICANT: Tumas, Daniel

APPLICANT: Watanabe, Colin K.

APPLICANT: Williams, P. Mickey

APPLICANT: Wood, William I.

APPLICANT: Zhang, Zemin

TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic

TITLE OF INVENTION: Acids Encoding the Same

; PRIOR APPLICATION NUMBER: 60/090695
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090696
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090862
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/090863
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/091360
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091478
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091544
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091626
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091633
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091978
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/092182
; PRIOR FILING DATE: 1998-07-09

Query Match 2.2%; Score 33.2; DB 9; Length 3226;
Best Local Similarity 54.0%; Pred. No. 11;
Matches 68; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY 1355 GTCTCTGAAATATGTTGGCCAGGCGCTGTGTGTCACCTGTGGAGAGCAGGACT 1414
Db 1998 GTCTATTACATGATTTTCATCTGTCAGCATTTGGGAAGCTGTCCATTGAAAGGCCT 2057
QY 1415 GGGATGGAATAAATATGTTGTACAGGAACGCCAATATACCATCACCAAGTGCCTGGA 1474
Db 2058 TGGATTATCCCTGTACTTGAACATGAACACTGAATATTATGCCCGTGTTCAGAGTTTGA 2117
QY 1475 AGGAGC 1480
Db 2118 ATGAGC 2123

RESULT 32
US-09-989-734-352
; Sequence 352, Application US/09989734
; Publication No. US20030003531A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2730PIC64
; CURRENT FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/075945
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/087106
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087609
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087827
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: 60/088021
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088025
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088026
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088028
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088029
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088030
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088033
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088326
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088167
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088202
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088212
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088217
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088655
; PRIOR FILING DATE: 1998-06-09
; PRIOR APPLICATION NUMBER: 60/088734
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088738
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088742
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088810
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088824
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088826
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088858
; PRIOR FILING DATE: 1998-06-11

;	PRIOR FILING DATE:	1998-06-25
;	PRIOR APPLICATION NUMBER:	60/090695
;	PRIOR FILING DATE:	1998-06-25
;	PRIOR APPLICATION NUMBER:	60/090696
;	PRIOR FILING DATE:	1998-06-25
;	PRIOR APPLICATION NUMBER:	60/090862
;	PRIOR FILING DATE:	1998-06-26
;	PRIOR APPLICATION NUMBER:	60/090863
;	PRIOR FILING DATE:	1998-06-26
;	PRIOR APPLICATION NUMBER:	60/091360
;	PRIOR FILING DATE:	1998-07-01
;	PRIOR APPLICATION NUMBER:	60/091478
;	PRIOR FILING DATE:	1998-07-02
;	PRIOR APPLICATION NUMBER:	60/091544
;	PRIOR FILING DATE:	1998-07-01
;	PRIOR APPLICATION NUMBER:	60/091519
;	PRIOR FILING DATE:	1998-07-02
;	PRIOR APPLICATION NUMBER:	60/091626
;	PRIOR FILING DATE:	1998-07-02
;	PRIOR APPLICATION NUMBER:	60/091633
;	PRIOR FILING DATE:	1998-07-02
;	PRIOR APPLICATION NUMBER:	60/091978
;	PRIOR FILING DATE:	1998-07-07
;	PRIOR APPLICATION NUMBER:	60/091982
;	PRIOR FILING DATE:	1998-07-07
;	PRIOR APPLICATION NUMBER:	60/092182
;	PRIOR FILING DATE:	1998-07-09
;		
Query Match		
Best Local Similarity 2.2%; Score 33.2.2.		
Matches 68; Conservative 0; Mismatched		
Oy	1355	GTCCTCGTGGAAATGATTGGCGCAGGCCGCCC
Db	1998	GTCTCATTTACCAATGCATTTCAGTCGTGTCGA
Oy	1415	GGGATGGAAAAATAAATATGTTGTCACGAGGA
Db	2058	TGGATTATCCCTGTACTTGTAACACATGA
Oy	1475	AGGAGC 1480
Db	2118	ATGAGC 2123
RESULT 33		
US-10-028-072-463		
;	Sequence 463, Application US/10028072	
;	Publication No. US20030004311AI	
;	GENERAL INFORMATION:	
;	APPLICANT: Baker, Kevin P.	
;	APPLICANT: Beresini, Maureen	
;	APPLICANT: DeForge, Laura	
;	APPLICANT: Desnoyers, Luc	
;	APPLICANT: Filvaroff, Ellen	
;	APPLICANT: Gao, Wei-Qiang	
;	APPLICANT: Gerritsen, Mary E.	
;	APPLICANT: Goddard, Audrey	
;	APPLICANT: Godowski, Paul J.	
;	APPLICANT: Gurney, Austin L.	
;	APPLICANT: Sherwood, Steven	
;	APPLICANT: Smith, Victoria	
;	APPLICANT: Stewart, Timothy A.	
;	APPLICANT: Tumas, Daniel	
;	APPLICANT: Watanabe, Collin K	
;	APPLICANT: Wood, William	
;	APPLICANT: Zhang	
;	TITLE OF INVENTION:	
;	FILE REFERENCE:	
;	CURRENT APPLICATION NUMBER: US/10/028, 072	
;	CURRENT FILING DATE: 2001-12-19	
;	PRIOR APPLICATION NUMBER: 60/049911	
;	PRIOR FILING DATE: 1997-06-18	

7	PRIOR FILING DATE: 1997-12-11	
7	PRIOR APPLICATION NUMBER: 60/069334	
7	PRIOR FILING DATE: 1997-12-11	
7	PRIOR APPLICATION NUMBER: 60/069694	
7	PRIOR FILING DATE: 1997-12-16	
7	PRIOR APPLICATION NUMBER: 60/072320	
7	PRIOR FILING DATE: 1998-01-23	
7	PRIOR APPLICATION NUMBER: 60/073612	
7	PRIOR FILING DATE: 1998-02-04	
7	PRIOR APPLICATION NUMBER: 60/074086	
7	PRIOR FILING DATE: 1998-02-09	
7	PRIOR APPLICATION NUMBER: 60/074092	
7	PRIOR FILING DATE: 1998-02-09	
7	PRIOR APPLICATION NUMBER: 60/077791	
7	PRIOR FILING DATE: 1998-03-12	
7	PRIOR APPLICATION NUMBER: 60/078910	
7	PRIOR FILING DATE: 1998-03-20	
7	PRIOR APPLICATION NUMBER: 60/079294	
7	PRIOR FILING DATE: 1998-03-25	
7	PRIOR APPLICATION NUMBER: 60/079663	
7	PRIOR FILING DATE: 1998-02-27	
7	PRIOR APPLICATION NUMBER: 60/079728	
7	PRIOR FILING DATE: 1998-03-27	
7	PRIOR APPLICATION NUMBER: 60/080165	
7	PRIOR FILING DATE: 1998-03-31	
7	PRIOR APPLICATION NUMBER: 60/081203	
7	PRIOR FILING DATE: 1998-04-09	
7	PRIOR APPLICATION NUMBER: 60/081229	
7	PRIOR FILING DATE: 1998-04-09	
7	PRIOR APPLICATION NUMBER: 60/081695	
7	PRIOR FILING DATE: 1998-04-14	
7	PRIOR APPLICATION NUMBER: 60/081817	
7	PRIOR FILING DATE: 1998-04-15	
7	PRIOR APPLICATION NUMBER: 60/081818	
7	PRIOR FILING DATE: 1998-04-15	
7	PRIOR APPLICATION NUMBER: 60/082999	
7	PRIOR FILING DATE: 1998-04-24	
7	PRIOR APPLICATION NUMBER: 60/083322	
7	PRIOR FILING DATE: 1998-04-28	
7	PRIOR APPLICATION NUMBER: 60/083545	
7	PRIOR FILING DATE: 1998-04-29	
7	PRIOR APPLICATION NUMBER: 60/084600	
7	PRIOR FILING DATE: 1998-05-07	
7	PRIOR APPLICATION NUMBER: 60/084627	
7	PRIOR FILING DATE: 1998-05-07	
7	PRIOR APPLICATION NUMBER: 60/084637	
7	PRIOR FILING DATE: 1998-05-07	
7	PRIOR APPLICATION NUMBER: 60/085149	
7	PRIOR FILING DATE: 1998-05-12	
7	PRIOR APPLICATION NUMBER: 60/085323	
7	PRIOR FILING DATE: 1998-05-13	
7	PRIOR APPLICATION NUMBER: 60/085338	
7	PRIOR FILING DATE: 1998-05-13	
7	PRIOR APPLICATION NUMBER: 60/085339	
7	PRIOR FILING DATE: 1998-05-13	
7	PRIOR APPLICATION NUMBER: 60/085579	
7	PRIOR FILING DATE: 1998-05-15	
7	PRIOR APPLICATION NUMBER: 60/085697	
7	PRIOR FILING DATE: 1998-05-15	
7	PRIOR APPLICATION NUMBER: 60/085704	
7	PRIOR FILING DATE: 1998-05-15	
7	PRIOR APPLICATION NUMBER: 60/086414	
7	PRIOR FILING DATE: 1998-05-22	
7	PRIOR APPLICATION NUMBER: 60/086430	
7	PRIOR FILING DATE: 1998-05-22	
7	PRIOR APPLICATION NUMBER: 60/087106	
7	PRIOR FILING DATE: 1998-05-28	
7	PRIOR APPLICATION NUMBER: 60/088026	
7	PRIOR FILING DATE: 1998-06-04	
7	PRIOR APPLICATION NUMBER: 60/088730	
7	PRIOR FILING DATE: 1998-06-10	
7	PRIOR APPLICATION NUMBER: 60/088741	
7	PRIOR FILING DATE: 1998-06-10	

;; PRIOR APPLICATION NUMBER: 60/088810
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088858
;; PRIOR FILING DATE: 19/98-06-11
;; PRIOR APPLICATION NUMBER: 60/089532
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089599
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089907
;; PRIOR FILING DATE: 1998-06-18
;; PRIOR APPLICATION NUMBER: 60/089947
;; PRIOR FILING DATE: 1998-06-19
;; PRIOR APPLICATION NUMBER: 60/090349
;; PRIOR FILING DATE: 1998-06-23
;; PRIOR APPLICATION NUMBER: 60/090429
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090445
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090538
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090863
;; PRIOR FILING DATE: 1998-06-26
;; PRIOR APPLICATION NUMBER: 60/091360
;; PRIOR FILING DATE: 1998-07-01
;; PRIOR APPLICATION NUMBER: 60/091519
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091982
;; PRIOR FILING DATE: 1998-07-07

Query Match 2.2%; Score 33.2; DB 9; Length 3226;
Best Local Similarity 54.0%; Pred No. 11;

Matches	68;	Conservative	0;	Mismatches	58;	Indels	0;	Gaps	0;
QY	1355	GTCTCTGAAATATGTTGGCGCAGGCGCTGATGTCACCTGTGGAGGAGGAGGACT	1414						
Db	1998	GTCTCATTAACTGATTCATGCTGTCAGCTTGGGAAGCTGTCCATTGAAAGGCT	2057						
QY	1415	GGATGGGAAATAATGTGTACAGGAACGGCAATATACCATCACCATGCGCTGGA	1474						
Db	2058	TGATTTATCCCTGACTTGTAACATGAATGAAATATTGCGCGTGTTCGAAGGTTGA	2117						
QY	1475	AGGAGC 1480							
Db	2118	ATGAGC 2123							

RESULT 34

US-09-997-653-352
; Sequence 352, Application US/09997653
; Publication No. US2003008297A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey

;; APPLICANT: Wood, William I.
;; APPLICANT: Zhang, Zemin
;; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
;; FILE REFERENCE: P2730PIC38
;; CURRENT APPLICATION NUMBER: US/09/997,653
;; CURRENT FILING DATE: 2001-11-15
;; PRIOR APPLICATION NUMBER: 60/049787
;; PRIOR FILING DATE: 1997-06-16
;; PRIOR APPLICATION NUMBER: 60/062250
;; PRIOR FILING DATE: 1997-10-17
;; PRIOR APPLICATION NUMBER: 60/065186
;; PRIOR FILING DATE: 1997-11-12
;; PRIOR APPLICATION NUMBER: 60/065311
;; PRIOR FILING DATE: 1997-11-13
;; PRIOR APPLICATION NUMBER: 60/066770
;; PRIOR FILING DATE: 1997-11-24
;; PRIOR APPLICATION NUMBER: 60/075945
;; PRIOR FILING DATE: 1998-02-25
;; PRIOR APPLICATION NUMBER: 60/078910
;; PRIOR FILING DATE: 1998-03-20
;; PRIOR APPLICATION NUMBER: 60/083322
;; PRIOR FILING DATE: 1998-04-28
;; PRIOR APPLICATION NUMBER: 60/084600
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/087106
;; PRIOR FILING DATE: 1998-05-28
;; PRIOR APPLICATION NUMBER: 60/087607
;; PRIOR FILING DATE: 1998-06-02
;; PRIOR APPLICATION NUMBER: 60/087609
;; PRIOR FILING DATE: 1998-06-02
;; PRIOR APPLICATION NUMBER: 60/087759
;; PRIOR FILING DATE: 1998-06-02
;; PRIOR APPLICATION NUMBER: 60/087827
;; PRIOR FILING DATE: 1998-06-03
;; PRIOR APPLICATION NUMBER: 60/088021
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088025
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088026
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088028
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088029
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088030
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088033
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088326
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088167
;; PRIOR FILING DATE: 1998-06-05
;; PRIOR APPLICATION NUMBER: 60/088202
;; PRIOR FILING DATE: 1998-06-05
;; PRIOR APPLICATION NUMBER: 60/088212
;; PRIOR FILING DATE: 1998-06-05
;; PRIOR APPLICATION NUMBER: 60/088217
;; PRIOR FILING DATE: 1998-06-05
;; PRIOR APPLICATION NUMBER: 60/088655
;; PRIOR FILING DATE: 1998-06-09
;; PRIOR APPLICATION NUMBER: 60/088734
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088738
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088742
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088810
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088824
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088826
;; PRIOR FILING DATE: 1998-06-10

	PRIOR APPLICATION NUMBER:	60/088958
	PRIOR FILING DATE:	1998-06-11
	PRIOR APPLICATION NUMBER:	60/088861
	PRIOR FILING DATE:	1998-06-11
	PRIOR APPLICATION NUMBER:	60/088876
	PRIOR FILING DATE:	1998-06-11
	PRIOR APPLICATION NUMBER:	60/089105
	PRIOR FILING DATE:	1998-06-12
	PRIOR APPLICATION NUMBER:	60/089440
	PRIOR FILING DATE:	1998-06-16
	PRIOR APPLICATION NUMBER:	60/089512
	PRIOR FILING DATE:	1998-06-16
	PRIOR APPLICATION NUMBER:	60/089514
	PRIOR FILING DATE:	1998-06-16
	PRIOR APPLICATION NUMBER:	60/089532
	PRIOR FILING DATE:	1998-06-17
	PRIOR APPLICATION NUMBER:	60/089538
	PRIOR FILING DATE:	1998-06-17
	PRIOR APPLICATION NUMBER:	60/089598
	PRIOR FILING DATE:	1998-06-17
	PRIOR APPLICATION NUMBER:	60/089599
	PRIOR FILING DATE:	1998-06-17
	PRIOR APPLICATION NUMBER:	60/089600
	PRIOR FILING DATE:	1998-06-17
	PRIOR APPLICATION NUMBER:	60/089653
	PRIOR FILING DATE:	1998-06-17
	PRIOR APPLICATION NUMBER:	60/089801
	PRIOR FILING DATE:	1998-06-18
	PRIOR APPLICATION NUMBER:	60/089907
	PRIOR FILING DATE:	1998-06-18
	PRIOR APPLICATION NUMBER:	60/089908
	PRIOR FILING DATE:	1998-06-18
	PRIOR APPLICATION NUMBER:	60/089947
	PRIOR FILING DATE:	1998-06-19
	PRIOR APPLICATION NUMBER:	60/089948
	PRIOR FILING DATE:	1998-06-19
	PRIOR APPLICATION NUMBER:	60/089952
	PRIOR FILING DATE:	1998-06-19
	PRIOR APPLICATION NUMBER:	60/090246
	PRIOR FILING DATE:	1998-06-22
	PRIOR APPLICATION NUMBER:	60/090252
	PRIOR FILING DATE:	1998-06-22
	PRIOR APPLICATION NUMBER:	60/090254
	PRIOR FILING DATE:	1998-06-22
	PRIOR APPLICATION NUMBER:	60/090349
	PRIOR FILING DATE:	1998-06-23
	PRIOR APPLICATION NUMBER:	60/090355
	PRIOR FILING DATE:	1998-06-23
	PRIOR APPLICATION NUMBER:	60/090429
	PRIOR FILING DATE:	1998-06-24
	PRIOR APPLICATION NUMBER:	60/090431
	PRIOR FILING DATE:	1998-06-24
	PRIOR APPLICATION NUMBER:	60/090435
	PRIOR FILING DATE:	1998-06-24
	PRIOR APPLICATION NUMBER:	60/090444
	PRIOR FILING DATE:	1998-06-24
	PRIOR APPLICATION NUMBER:	60/090445
	PRIOR FILING DATE:	1998-06-24
	PRIOR APPLICATION NUMBER:	60/090540
	PRIOR FILING DATE:	1998-06-24
	PRIOR APPLICATION NUMBER:	60/090542
	PRIOR FILING DATE:	1998-06-24
	PRIOR APPLICATION NUMBER:	60/090557
	PRIOR FILING DATE:	1998-06-24
	PRIOR APPLICATION NUMBER:	60/090676
	PRIOR FILING DATE:	1998-06-25
	PRIOR APPLICATION NUMBER:	60/090678
	PRIOR FILING DATE:	1998-06-25
	PRIOR APPLICATION NUMBER:	60/090690

; PRIOR FILING DATE: 1998-06-25
 ; PRIOR APPLICATION NUMBER: 60/090694
 ; PRIOR FILING DATE: 1998-06-25
 ; PRIOR APPLICATION NUMBER: 60/090695
 ; PRIOR FILING DATE: 1998-06-25
 ; PRIOR APPLICATION NUMBER: 60/090696
 ; PRIOR FILING DATE: 1998-06-25
 ; PRIOR APPLICATION NUMBER: 60/090862
 ; PRIOR FILING DATE: 1998-06-26
 ; PRIOR APPLICATION NUMBER: 60/090863
 ; PRIOR FILING DATE: 1998-06-26
 ; PRIOR APPLICATION NUMBER: 60/091360
 ; PRIOR FILING DATE: 1998-07-01
 ; PRIOR APPLICATION NUMBER: 60/091478
 ; PRIOR FILING DATE: 1998-07-02
 ; PRIOR APPLICATION NUMBER: 60/091544
 ; PRIOR FILING DATE: 1998-07-01
 ; PRIOR APPLICATION NUMBER: 60/091519
 ; PRIOR FILING DATE: 1998-07-02
 ; PRIOR APPLICATION NUMBER: 60/091626
 ; PRIOR FILING DATE: 1998-07-02
 ; PRIOR APPLICATION NUMBER: 60/091633
 ; PRIOR FILING DATE: 1998-07-02
 ; PRIOR APPLICATION NUMBER: 60/091978
 ; PRIOR FILING DATE: 1998-07-07
 ; PRIOR APPLICATION NUMBER: 60/091982
 ; PRIOR FILING DATE: 1998-07-07
 ; PRIOR APPLICATION NUMBER: 60/092182
 ; PRIOR FILING DATE: 1998-07-09

```

RESULT 35
; Sequence 352, Application US/09993667
; Publication No. US20030002187A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Denoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kijavlin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.

```

APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P2730PlC4
CURRENT FILING DATE: 2001-11-14
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/049787
PRIOR FILING DATE: 1997-06-16
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/065186
PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066770
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/075945
PRIOR FILING DATE: 1998-02-25
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/087106
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087607
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087609
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087759
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087827
PRIOR FILING DATE: 1998-06-03
PRIOR APPLICATION NUMBER: 60/088021
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088025
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088026
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088028
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088029
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088030
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088033
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088326
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088167
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088202
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088212
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088217
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088655
PRIOR FILING DATE: 1998-06-09
PRIOR APPLICATION NUMBER: 60/088734
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088738
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088742
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088810
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088824
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088826
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088858
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088861
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088876
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/089105
PRIOR FILING DATE: 1998-06-12
PRIOR APPLICATION NUMBER: 60/089440
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089512
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089514
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089532
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089538
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089598
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089599
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089600
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089653
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089801
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089907
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089908
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089947
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089948
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089952
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/090246
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090252
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090254
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090349
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090355
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090429
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090431
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090435
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090444
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090445
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090472
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090535
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090540
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090542
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090557
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090676
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090678
PRIOR FILING DATE: 1998-06-25

```

RESULT 36
US-10-121-049-463
; Sequence 463, Application US/10121049
; Publication No. US20030022339A1
;
; GENERAL INFORMATION:
;
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
;
; TITLE OF INVENTION: SECRETED AND TRA
;
; TITLE OF INVENTION: ACIDS ENCODING
;
; FILE REFERENCE: P3330R1C17

```

	Query Match	2.2%	Score 33.2	DB 9	Length 3236
	Best Local Similarity	54.0%	Pred. No. 11		
	Matches 68	Conservative 0	Mismatches 58	Indels 0	Gaps 0
QY	1355	GTCTCCTGAAAAATGTAATTCGCGCAGGGCCTGTAGTGGTGACCTCTGTGGAGGACGACGGACT	1414		
Db	1998	GTCTCATTTAAACAATGCATTTTCAGCTCGTCGACATTTGGGAAAGCTGTCCATTGGAAGGCGCT	2057		
QY	1415	GGGATGGAAAAATTAATGTGGTGTACAGGAAACGGCAAAATATACCAATACCAAGTGGCCCTGGA	1474		
Db	2058	TGGATTTATCCCTGTACTTTGAAACACATGAACCTGAAATTTATGCCCGTGTCTTCAAGGTTTGA	2117		

APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Kljavin, Ivar J.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K.
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P2730P1C18
CURRENT APPLICATION NUMBER: US/09/990,562
CURRENT FILING DATE: 2001-11-14
PRIOR FILING DATE: 1997-06-16
PRIOR APPLICATION NUMBER: 60/049787
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/065186
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/066770
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/075945
PRIOR FILING DATE: 1998-02-25
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/087106
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087607
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087609
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087759
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087827
PRIOR FILING DATE: 1998-06-03
PRIOR APPLICATION NUMBER: 60/088021
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088025
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088026
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088028
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088029
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088030
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088033
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088326
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088167
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088202
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088212
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088217
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088655

PRIOR FILING DATE: 1998-06-09
PRIOR APPLICATION NUMBER: 60/088734
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088738
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088742
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088810
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088824
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088826
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088858
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088861
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088876
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/089105
PRIOR FILING DATE: 1998-06-12
PRIOR APPLICATION NUMBER: 60/089440
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089512
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089514
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089532
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089538
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089598
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089599
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089600
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089653
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089801
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089907
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089908
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089947
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089948
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089952
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/090246
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090252
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090254
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090349
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090355
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090429
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090431
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090435
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090444
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090445
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090472
PRIOR FILING DATE: 1998-06-24

;; PRIOR APPLICATION NUMBER: 60/090535
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090540
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090542
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090557
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090676
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090678
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090690
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090694
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090695
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090696
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090862
;; PRIOR FILING DATE: 1998-06-26
;; PRIOR APPLICATION NUMBER: 60/090863
;; PRIOR FILING DATE: 1998-06-26
;; PRIOR APPLICATION NUMBER: 60/091360
;; PRIOR FILING DATE: 1998-07-01
;; PRIOR APPLICATION NUMBER: 60/091478
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091544
;; PRIOR FILING DATE: 1998-07-01
;; PRIOR APPLICATION NUMBER: 60/091519
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091626
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091633
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091978
;; PRIOR FILING DATE: 1998-07-07
;; PRIOR APPLICATION NUMBER: 60/091982
;; PRIOR FILING DATE: 1998-07-07
;; PRIOR APPLICATION NUMBER: 60/092182
;; PRIOR FILING DATE: 1998-07-09

Query Match 2.28; Score 33.2; DB 9; Length 3226;

Best Local Similarity 54.08; Pred. No. 11;

Matches 68; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

Qy 1355 GTCTCTGAAATGTTATGGCGCAGGCGCTGATGGTGCACCTGTGGAGGAGGAGCT 1414

Db 1998 GTCTATTACATGCAITTCAGCTCTCAGCATTTGGAAAGCTGTCCATTGAAAGGCCT 2057

Qy 1415 GGGATGGAATAAATGTGTGACAGGAAACGGCAAAATATACCATCACCAGTGGCCCTGGA 1474

Db 2058 TGGATTATCCCTGACTTGAACATGAACATGAATATGCCCGTCTTCAAGTTTGA 2117

Qy 1475 AGGAGC 1480

Db 2118 ATGAGC 2123

RESULT 41

US-09-997-428-352

Sequence 352, Application US/09997428

Publication No. US20030027162A1

GENERAL INFORMATION:

APPLICANT: Ashkenazi, Avi J.

APPLICANT: Baker, Kevin P.

APPLICANT: Botstein, David

APPLICANT: Desnoyers, Luc

APPLICANT: Eaton, Dan L.

APPLICANT: Ferrara, Napoleone

APPLICANT: Fong, Sherman

APPLICANT: Gerber, Hanspeter

;; APPLICANT: Gerritsen, Mary E.
;; APPLICANT: Goddard, Audrey
;; APPLICANT: Godowski, Paul J.
;; APPLICANT: Grimaldi, J. Christopher
;; APPLICANT: Gurney, Austin L.
;; APPLICANT: Kljavin, Ivar J.
;; APPLICANT: Napier, Mary A.
;; APPLICANT: Pan, James
;; APPLICANT: Paoni, Nicholas F.
;; APPLICANT: Roy, Margaret Ann
;; APPLICANT: Stewart, Timothy A.
;; APPLICANT: Tumas, Daniel
;; APPLICANT: Watanabe, Colin K.
;; APPLICANT: Williams, P. Mickey
;; APPLICANT: Wood, William I.
;; APPLICANT: Zhang, Zemin
;; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
;; TITLE OF INVENTION: Acids Encoding the Same
;; FILE REFERENCE: P2730P1C44
;; CURRENT APPLICATION NUMBER: US/09/997,428
;; CURRENT FILING DATE: 2001-11-15
;; PRIOR APPLICATION NUMBER: 60/049787
;; PRIOR FILING DATE: 1997-06-16
;; PRIOR APPLICATION NUMBER: 60/062250
;; PRIOR FILING DATE: 1997-10-17
;; PRIOR APPLICATION NUMBER: 60/065186
;; PRIOR FILING DATE: 1997-11-12
;; PRIOR APPLICATION NUMBER: 60/065311
;; PRIOR FILING DATE: 1997-11-13
;; PRIOR APPLICATION NUMBER: 60/066770
;; PRIOR FILING DATE: 1997-11-24
;; PRIOR APPLICATION NUMBER: 60/075945
;; PRIOR FILING DATE: 1998-02-25
;; PRIOR APPLICATION NUMBER: 60/078910
;; PRIOR FILING DATE: 1998-03-20
;; PRIOR APPLICATION NUMBER: 60/083322
;; PRIOR FILING DATE: 1998-04-28
;; PRIOR APPLICATION NUMBER: 60/084600
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/087106
;; PRIOR FILING DATE: 1998-05-28
;; PRIOR APPLICATION NUMBER: 60/087607
;; PRIOR FILING DATE: 1998-06-02
;; PRIOR APPLICATION NUMBER: 60/087609
;; PRIOR FILING DATE: 1998-06-02
;; PRIOR APPLICATION NUMBER: 60/087759
;; PRIOR FILING DATE: 1998-06-02
;; PRIOR APPLICATION NUMBER: 60/087827
;; PRIOR FILING DATE: 1998-06-03
;; PRIOR APPLICATION NUMBER: 60/088021
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088025
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088026
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088028
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088029
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088030
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088033
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088326
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088167
;; PRIOR FILING DATE: 1998-06-05
;; PRIOR APPLICATION NUMBER: 60/088202
;; PRIOR FILING DATE: 1998-06-05
;; PRIOR APPLICATION NUMBER: 60/088212
;; PRIOR FILING DATE: 1998-06-05
;; PRIOR APPLICATION NUMBER: 60/088217
;; PRIOR FILING DATE: 1998-06-05

;; PRIOR APPLICATION NUMBER: 60/088655
;; PRIOR FILING DATE: 1998-06-09
;; PRIOR APPLICATION NUMBER: 60/088734
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088738
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088742
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088810
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088824
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088826
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088858
;; PRIOR FILING DATE: 1998-06-11
;; PRIOR APPLICATION NUMBER: 60/088861
;; PRIOR FILING DATE: 1998-06-11
;; PRIOR APPLICATION NUMBER: 60/088876
;; PRIOR FILING DATE: 1998-06-11
;; PRIOR APPLICATION NUMBER: 60/089105
;; PRIOR FILING DATE: 1998-06-12
;; PRIOR APPLICATION NUMBER: 60/089440
;; PRIOR FILING DATE: 1998-06-16
;; PRIOR APPLICATION NUMBER: 60/089512
;; PRIOR FILING DATE: 1998-06-16
;; PRIOR APPLICATION NUMBER: 60/089514
;; PRIOR FILING DATE: 1998-06-16
;; PRIOR APPLICATION NUMBER: 60/089532
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089538
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089598
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089599
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089600
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089653
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089801
;; PRIOR FILING DATE: 1998-06-18
;; PRIOR APPLICATION NUMBER: 60/089907
;; PRIOR FILING DATE: 1998-06-18
;; PRIOR APPLICATION NUMBER: 60/089908
;; PRIOR FILING DATE: 1998-06-18
;; PRIOR APPLICATION NUMBER: 60/089947
;; PRIOR FILING DATE: 1998-06-19
;; PRIOR APPLICATION NUMBER: 60/089948
;; PRIOR FILING DATE: 1998-06-19
;; PRIOR APPLICATION NUMBER: 60/089952
;; PRIOR FILING DATE: 1998-06-19
;; PRIOR APPLICATION NUMBER: 60/090246
;; PRIOR FILING DATE: 1998-06-22
;; PRIOR APPLICATION NUMBER: 60/090252
;; PRIOR FILING DATE: 1998-06-22
;; PRIOR APPLICATION NUMBER: 60/090254
;; PRIOR FILING DATE: 1998-06-22
;; PRIOR APPLICATION NUMBER: 60/090349
;; PRIOR FILING DATE: 1998-06-23
;; PRIOR APPLICATION NUMBER: 60/090355
;; PRIOR FILING DATE: 1998-06-23
;; PRIOR APPLICATION NUMBER: 60/090429
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090431
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090435
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090444
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090445
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090472

;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090535
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090540
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090542
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090557
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090676
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090678
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090690
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090694
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090695
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090696
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090862
;; PRIOR FILING DATE: 1998-06-26
;; PRIOR APPLICATION NUMBER: 60/090863
;; PRIOR FILING DATE: 1998-06-26
;; PRIOR APPLICATION NUMBER: 60/091360
;; PRIOR FILING DATE: 1998-07-01
;; PRIOR APPLICATION NUMBER: 60/091478
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091544
;; PRIOR FILING DATE: 1998-07-01
;; PRIOR APPLICATION NUMBER: 60/091519
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091626
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091633
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091978
;; PRIOR FILING DATE: 1998-07-07
;; PRIOR APPLICATION NUMBER: 60/091982
;; PRIOR FILING DATE: 1998-07-07
;; PRIOR APPLICATION NUMBER: 60/092182
;; PRIOR FILING DATE: 1998-07-09

Query Match 2.2%; Score 33.2; DB 9; Length 3226;
Best Local Similarity 54.0%; Pred. No. 11;
Matches 68; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

Qy 1355 GTCTCTGAAAAATGTTATGCGCGAGGCGCTGTGTTGTCACCTGTGTGAGGAGGAGCT 1414
Db 1998 GTCTCATTAACAATGCATTTTCAGCTCGTCAGCATTTGGGAAGCTGTCCATTGAAAAAGCCT 2057
Qy 1415 GGGATGAAAAATAAATGTGTACAGGAACGCAATATACCATCACCAGTGCCTGGA 1474
Db 2058 TGGATTATTCCTGCTACTTGAACATGAAATGAAATATGCGCGTGTTCAGAGTTTGA 2117
Qy 1475 AGGAGC 1480
Db 2118 ATGAGC 2123

RESULT 42

US-09-997-666-352
; Sequence 352, Application US/09997666
; Publication No. US20030027163A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman

APPLICANT: Gerber, Hanspeter
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Kljavin, Ivar J.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K.
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P2730PIC42
CURRENT APPLICATION NUMBER: US/09/997,666
CURRENT FILING DATE: 2001-11-15
PRIOR APPLICATION NUMBER: 60/049787
PRIOR FILING DATE: 1997-06-16
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/065186
PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066770
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/075945
PRIOR FILING DATE: 1998-02-25
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/087106
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087607
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087759
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087827
PRIOR FILING DATE: 1998-06-03
PRIOR APPLICATION NUMBER: 60/088021
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088025
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088026
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088028
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088029
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088030
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088033
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088326
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088167
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088202
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088212
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088217
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088655
PRIOR FILING DATE: 1998-06-09
PRIOR APPLICATION NUMBER: 60/088734
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088738
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088742
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088810
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088824
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088826
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088858
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088861
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088876
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/089105
PRIOR FILING DATE: 1998-06-12
PRIOR APPLICATION NUMBER: 60/089440
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089512
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089514
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089532
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089538
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089598
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089599
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089600
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089653
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089801
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089907
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089908
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089947
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089948
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089952
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/090246
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090252
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090254
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090349
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090355
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090429
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090431
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090435
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090444
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090445
PRIOR FILING DATE: 1998-06-24

```

; PRIOR APPLICATION NUMBER: 60/090472
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090535
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090540
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090542
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090676
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090678
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090690
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090694
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090695
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090696
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090862
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/090863
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/091360
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091478
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091544
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091626
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091633
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091978
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/092182
; PRIOR FILING DATE: 1998-07-09

```

```

Query Match      2.2%  Score 33.2; DB 9; Length 3226;
Best Local Similarity 54.0%; Pred. No. 11;
Matches 68; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY 1355 GTCCTCGAAAAATGATTTGGCGAGGCGCTGATGGTGCACCTGTGGAGGAGGAGGACT 1414
      ||||| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1998 GTCTCATTAACAATGCAATTCAGCTGCTCAGCATTTGGGAAGCTGTCCATTGAAAAGGCCT 2057

QY 1415 GGGATGGAAAAATAAATGTGTGTACAGAAACGGCAATATACCATCACCAGTGGCCTGGA 1474
      ||||| || || || || || || || || || || || || || || || || || || ||
Db 2058 TGGATTATCCCTGCTACTTGAACATGAAACTGAAATATATGCCCCGTTTCAAGGTTGA 2117

QY 1475 AGGAGC 1480
      |||||
Db 2118 ATGAGC 2123

```

```

RESULT 43
US-10-175-746-463
; Sequence 463, Application US/10175746
; Publication No. US20030027270A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen

```

```

; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C353
; CURRENT APPLICATION NUMBER: US/10/175,746
; CURRENT FILING DATE: 2002-06-19
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 463
; LENGTH: 3226
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-175-746-463

Query Match      2.2%  Score 33.2; DB 9; Length 3226;
Best Local Similarity 54.0%; Pred. No. 11;
Matches 68; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY 1355 GTCCTCGAAAAATGATTTGGCGAGGCGCTGATGGTGCACCTGTGGAGGAGGAGGACT 1414
      ||||| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1998 GTCTCATTAACAATGCAATTCAGCTGCTCAGCATTTGGGAAGCTGTCCATTGAAAAGGCCT 2057

QY 1415 GGGATGGAAAAATAAATGTGTGTACAGAAACGGCAATATACCATCACCAGTGGCCTGGA 1474
      ||||| || || || || || || || || || || || || || || || || || || ||
Db 2058 TGGATTATCCCTGCTACTTGAACATGAAACTGAAATATATGCCCCGTTTCAAGGTTGA 2117

QY 1475 AGGAGC 1480
      |||||
Db 2118 ATGAGC 2123

```

```

RESULT 44
US-10-176-918-463
; Sequence 463, Application US/10176918
; Publication No. US20030027275A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C382
; CURRENT APPLICATION NUMBER: US/10/176,918
; CURRENT FILING DATE: 2002-06-20
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 463
; LENGTH: 3226
; TYPE: DNA

```

ORGANISM: Homo Sapien
US-10-176-918-463

Query Match 2.28; Score 33.2; DB 9; Length 3226;
Best Local Similarity 54.0%; Pred. No. 11;
Matches 68; Conservative 0; Mismatches 58; Indels 0; Gaps 0;
Qy 1355 GTCTCCGAAATGATTGGCGCAGGCGCTGATGGTGACCTGTGGAGGACGAGGACT 1414
Db 1998 GTCTCATTAACAATGATTGAGCTGCTGACGATTTGGGAAGCTGTCCATTGAAAGGCGCT 2057
Qy 1415 GGGATGGAATAAATGTTGTACAGGAAGCGCAATATACCATCACCAGTGCCCTGGA 1474
Db 2058 TGGATTATCCCTGACTTGAACATGAACATGAATATATGCCGCTTTCAAGGTTGA 2117
Qy 1475 AGGAGC 1480
Db 2118 ATGAGC 2123

RESULT 45

US-10-176-921-463
Sequence 463, Application US/10176921
Publication No. US20030027276A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: DeForge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
TITLE OF INVENTION: ACIDS ENCODING THE SAME
FILE REFERENCE: P3330RIC288
CURRENT APPLICATION NUMBER: US/10/176.921
CURRENT FILING DATE: 2002-06-20
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 463
LENGTH: 3226
TYPE: DNA
ORGANISM: Homo Sapien
US-10-176-921-463
Query Match 2.28; Score 33.2; DB 9; Length 3226;
Best Local Similarity 54.0%; Pred. No. 11;
Matches 68; Conservative 0; Mismatches 58; Indels 0; Gaps 0;
Qy 1355 GTCTCCGAAATGATTGGCGCAGGCGCTGATGGTGACCTGTGGAGGACGAGGACT 1414
Db 1998 GTCTCATTAACAATGATTGAGCTGCTGACGATTTGGGAAGCTGTCCATTGAAAGGCGCT 2057
Qy 1415 GGGATGGAATAAATGTTGTACAGGAAGCGCAATATACCATCACCAGTGCCCTGGA 1474
Db 2058 TGGATTATCCCTGACTTGAACATGAACATGAATATATGCCGCTTTCAAGGTTGA 2117
Qy 1475 AGGAGC 1480
Db 2118 ATGAGC 2123

RESULT 46

US-09-990-711-352
Sequence 352, Application US/09990711
Publication No. US20030032023A1
GENERAL INFORMATION:
APPLICANT: Ashkenazi, Avi J.
APPLICANT: Baker, Kevin P.
APPLICANT: Botstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Eaton, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Fong, Sherman
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Kljavin, Ivar J.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K.
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
TITLE OF INVENTION: Acids Encoding the Same
FILE REFERENCE: P2730P1C2
CURRENT APPLICATION NUMBER: US/09/990.711
CURRENT FILING DATE: 2001-11-14
Prior Application Number: 60/049787
Prior Filing Date: 1997-06-16
Prior Application Number: 60/062250
Prior Filing Date: 1997-10-17
Prior Application Number: 60/065186
Prior Filing Date: 1997-11-12
Prior Application Number: 60/065311
Prior Filing Date: 1997-11-13
Prior Application Number: 60/066770
Prior Filing Date: 1997-11-24
Prior Application Number: 60/075945
Prior Filing Date: 1998-02-25
Prior Application Number: 60/078910
Prior Filing Date: 1998-03-20
Prior Application Number: 60/083322
Prior Filing Date: 1998-04-28
Prior Application Number: 60/084600
Prior Filing Date: 1998-05-07
Prior Application Number: 60/087106
Prior Filing Date: 1998-05-28
Prior Application Number: 60/087607
Prior Filing Date: 1998-06-02
Prior Application Number: 60/087609
Prior Filing Date: 1998-06-02
Prior Application Number: 60/087759
Prior Filing Date: 1998-06-02
Prior Application Number: 60/087827
Prior Filing Date: 1998-06-03
Prior Application Number: 60/088021
Prior Filing Date: 1998-06-04
Prior Application Number: 60/088025
Prior Filing Date: 1998-06-04
Prior Application Number: 60/088026
Prior Filing Date: 1998-06-04
Prior Application Number: 60/088028
Prior Filing Date: 1998-06-04
Prior Application Number: 60/088029
Prior Filing Date: 1998-06-04
Prior Application Number: 60/088030
Prior Filing Date: 1998-06-04

; PRIOR APPLICATION NUMBER: 60/088033
 ; PRIOR FILING DATE: 1998-06-04
 ; PRIOR APPLICATION NUMBER: 60/088326
 ; PRIOR FILING DATE: 1998-06-04
 ; PRIOR APPLICATION NUMBER: 60/088167
 ; PRIOR FILING DATE: 1998-06-05
 ; PRIOR APPLICATION NUMBER: 60/088202
 ; PRIOR FILING DATE: 1998-06-05
 ; PRIOR APPLICATION NUMBER: 60/088212
 ; PRIOR FILING DATE: 1998-06-05
 ; PRIOR APPLICATION NUMBER: 60/088217
 ; PRIOR FILING DATE: 1998-06-05
 ; PRIOR APPLICATION NUMBER: 60/088655
 ; PRIOR FILING DATE: 1998-06-09
 ; PRIOR APPLICATION NUMBER: 60/088734
 ; PRIOR FILING DATE: 1998-06-10
 ; PRIOR APPLICATION NUMBER: 60/088738
 ; PRIOR FILING DATE: 1998-06-10
 ; PRIOR APPLICATION NUMBER: 60/088742
 ; PRIOR FILING DATE: 1998-06-10
 ; PRIOR APPLICATION NUMBER: 60/088810
 ; PRIOR FILING DATE: 1998-06-10
 ; PRIOR APPLICATION NUMBER: 60/088824
 ; PRIOR FILING DATE: 1998-06-10
 ; PRIOR APPLICATION NUMBER: 60/088826
 ; PRIOR FILING DATE: 1998-06-10
 ; PRIOR APPLICATION NUMBER: 60/088858
 ; PRIOR FILING DATE: 1998-06-11
 ; PRIOR APPLICATION NUMBER: 60/088861
 ; PRIOR FILING DATE: 1998-06-11
 ; PRIOR APPLICATION NUMBER: 60/088876
 ; PRIOR FILING DATE: 1998-06-11
 ; PRIOR APPLICATION NUMBER: 60/089105
 ; PRIOR FILING DATE: 1998-06-12
 ; PRIOR APPLICATION NUMBER: 60/089440
 ; PRIOR FILING DATE: 1998-06-16
 ; PRIOR APPLICATION NUMBER: 60/089512
 ; PRIOR FILING DATE: 1998-06-16
 ; PRIOR APPLICATION NUMBER: 60/089514
 ; PRIOR FILING DATE: 1998-06-16
 ; PRIOR APPLICATION NUMBER: 60/089532
 ; PRIOR FILING DATE: 1998-06-17
 ; PRIOR APPLICATION NUMBER: 60/089538
 ; PRIOR FILING DATE: 1998-06-17
 ; PRIOR APPLICATION NUMBER: 60/089598
 ; PRIOR FILING DATE: 1998-06-17
 ; PRIOR APPLICATION NUMBER: 60/089599
 ; PRIOR FILING DATE: 1998-06-17
 ; PRIOR APPLICATION NUMBER: 60/089600
 ; PRIOR FILING DATE: 1998-06-17
 ; PRIOR APPLICATION NUMBER: 60/089653
 ; PRIOR FILING DATE: 1998-06-17
 ; PRIOR APPLICATION NUMBER: 60/089801
 ; PRIOR FILING DATE: 1998-06-18
 ; PRIOR APPLICATION NUMBER: 60/089907
 ; PRIOR FILING DATE: 1998-06-18
 ; PRIOR APPLICATION NUMBER: 60/089908
 ; PRIOR FILING DATE: 1998-06-18
 ; PRIOR APPLICATION NUMBER: 60/089947
 ; PRIOR FILING DATE: 1998-06-19
 ; PRIOR APPLICATION NUMBER: 60/089948
 ; PRIOR FILING DATE: 1998-06-19
 ; PRIOR APPLICATION NUMBER: 60/089952
 ; PRIOR FILING DATE: 1998-06-19
 ; PRIOR APPLICATION NUMBER: 60/090246
 ; PRIOR FILING DATE: 1998-06-22
 ; PRIOR APPLICATION NUMBER: 60/090252
 ; PRIOR FILING DATE: 1998-06-22
 ; PRIOR APPLICATION NUMBER: 60/090254
 ; PRIOR FILING DATE: 1998-06-22
 ; PRIOR APPLICATION NUMBER: 60/090349
 ; PRIOR FILING DATE: 1998-06-23
 ; PRIOR APPLICATION NUMBER: 60/090355

; PRIOR FILING DATE: 1998-06-23
 ; PRIOR APPLICATION NUMBER: 60/090429
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090431
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090435
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090444
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090445
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090472
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090535
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090540
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090542
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090557
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090676
 ; PRIOR FILING DATE: 1998-06-25
 ; PRIOR APPLICATION NUMBER: 60/090678
 ; PRIOR FILING DATE: 1998-06-25
 ; PRIOR APPLICATION NUMBER: 60/090690
 ; PRIOR FILING DATE: 1998-06-25
 ; PRIOR APPLICATION NUMBER: 60/090694
 ; PRIOR FILING DATE: 1998-06-25
 ; PRIOR APPLICATION NUMBER: 60/090695
 ; PRIOR FILING DATE: 1998-06-25
 ; PRIOR APPLICATION NUMBER: 60/090696
 ; PRIOR FILING DATE: 1998-06-25
 ; PRIOR APPLICATION NUMBER: 60/090862
 ; PRIOR FILING DATE: 1998-06-26
 ; PRIOR APPLICATION NUMBER: 60/090863
 ; PRIOR FILING DATE: 1998-06-26
 ; PRIOR APPLICATION NUMBER: 60/091360
 ; PRIOR FILING DATE: 1998-07-01
 ; PRIOR APPLICATION NUMBER: 60/091478
 ; PRIOR FILING DATE: 1998-07-02
 ; PRIOR APPLICATION NUMBER: 60/091544
 ; PRIOR FILING DATE: 1998-07-01
 ; PRIOR APPLICATION NUMBER: 60/091519
 ; PRIOR FILING DATE: 1998-07-02
 ; PRIOR APPLICATION NUMBER: 60/091626
 ; PRIOR FILING DATE: 1998-07-02
 ; PRIOR APPLICATION NUMBER: 60/091633
 ; PRIOR FILING DATE: 1998-07-02
 ; PRIOR APPLICATION NUMBER: 60/091978
 ; PRIOR FILING DATE: 1998-07-07
 ; PRIOR APPLICATION NUMBER: 60/091982
 ; PRIOR FILING DATE: 1998-07-07
 ; PRIOR APPLICATION NUMBER: 60/092182
 ; PRIOR FILING DATE: 1998-07-09

Query Match 2.2%; Score 33.2; DB 9; Length 3226;
 Best Local Similarity 54.0%; Pred. No. 11;
 Matches 68; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY 1355 GTCTCTGAAATGTTATGGCGCAGGCGCTGATGTCACCTGTGGAGGACGAGGACT 1414
 Db 1998 GTCTCATTPACATGTCATTTCAGTCGTCAGCATTTGGGAAGCTGTCATTGAAAGGCT 2057
 QY 1415 GGGATGGAAAAATAATGTGTACAGGAAACGGCAATATACATCACCAGTGGCCTGGA 1474
 Db 2058 TGGATTATCCCTGTACTTGAACATGAACATGAATATGCCCCGTGTTCAAGTTTGA 2117
 QY 1475 AGGAGC 1480
 Db 2118 ATGAGC 2123

RESULT 47

US-10-137-865-463
; Sequence 463, Application US/10137865
; Publication No. US20030032155A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C154
; CURRENT APPLICATION NUMBER: US/10/137,865
; CURRENT FILING DATE: 2002-05-03
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 463
; LENGTH: 3226
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-137-865-463

Query Match 2.2%; Score 33.2; DB 9; Length 3226;
Best Local Similarity 54.0%; Pred. No. 11;
Matches 68; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY	1355	GTCTCTGAAAAATGATTGGCGCAGGCGCTGTGACGAAACGGCAATATACCATCACCAGTGGCGCTGGA	1414
Db	1998	GTCTCATTAAACATGCAATTCAGCTCGTCAGCATTTGGGAAGCTGTCCATTGAAAAGGCCT	2057
QY	1415	GGATGAAAAATAATGTGTACAGCAACGCAATATACCATCACCAGTGGCGCTGGA	1474
Db	2058	TGGATTTATCCCTGTACTTGAACATGAAACTGAATTTATGCCCGTGTTCAGAGTTTGA	2117
QY	1475	AGGAGC 1480	
Db	2118	ATGAGC 2123	

RESULT 48

US-10-140-474-463
; Sequence 463, Application US/10140474
; Publication No. US20030032156A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William

; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C162
; CURRENT APPLICATION NUMBER: US/10/140,474
; CURRENT FILING DATE: 2002-05-06
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 463
; LENGTH: 3226
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-140-474-463

Query Match 2.2%; Score 33.2; DB 9; Length 3226;
Best Local Similarity 54.0%; Pred. No. 11;
Matches 68; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY	1355	GTCTCTGAAAAATGATTGGCGCAGGCGCTGTGACGAAACGGCAATATACCATCACCAGTGGCGCTGGA	1414
Db	1998	GTCTCATTAAACATGCAATTCAGCTCGTCAGCATTTGGGAAGCTGTCCATTGAAAAGGCCT	2057
QY	1415	GGATGAAAAATAATGTGTACAGCAACGCAATATACCATCACCAGTGGCGCTGGA	1474
Db	2058	TGGATTTATCCCTGTACTTGAACATGAAACTGAATTTATGCCCGTGTTCAGAGTTTGA	2117
QY	1475	AGGAGC 1480	
Db	2118	ATGAGC 2123	

RESULT 49

US-10-142-431-463
; Sequence 463, Application US/10142431
; Publication No. US20030036179A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C251
; CURRENT APPLICATION NUMBER: US/10/142,431
; CURRENT FILING DATE: 2002-05-10
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 463
; LENGTH: 3226
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-142-431-463

Query Match 2.2%; Score 33.2; DB 9; Length 3226;
Best Local Similarity 54.0%; Pred. No. 11;
Matches 68; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY	1355	GTCTCTGAAAAATGATTGGCGCAGGCGCTGTGACGAAACGGCAATATACCATCACCAGTGGCGCTGGA	1414
Db	1998	GTCTCATTAAACATGCAATTCAGCTCGTCAGCATTTGGGAAGCTGTCCATTGAAAAGGCCT	2057

QY 1415 GGGATGGAATAAATGCTGTACAGGAACGGCAAAATATACCATCACCAGTGGCCTGGA 1474
|||||
Db 2058 TGGATTATCCCTGCTACTTGAACATGAATTAATGCGCGTGTTCGAAGGTTGA 2117
|||||
QY 1475 AGGAGC 1480
|||||
Db 2118 ATGAGC 2123

RESULT 50.

US-10-143-114-463
; Sequence 463, Application US/10143114
; Publication No. US20030036180A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tunas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P330R1C211
; CURRENT APPLICATION NUMBER: US/10/143,114
; CURRENT FILING DATE: 2002-05-09
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 463
; LENGTH: 3226
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-143-114-463

Query Match 2.2%; Score 33.2; DB 9; Length 3226;
Best Local Similarity 54.0%; Pred. No. 11;
Matches 68; Conservative 0; Mismatches 58; Indels 0; Gaps 0;
QY 1355 GTCTCTGAAATAATGTTGGCGAGGCGCTGTGATGGTGACCTGTGGAGGAGCAGGACT 1414
|||||
Db 1998 GTCTCATTAAACAATGCAATTCAGCTCGTCAGCATTTGGGAAGCTGTCCATTGAAAGGCCT 2057
|||||
QY 1415 GGGATGGAATAAATGCTGTACAGGAACGGCAAAATATACCATCACCAGTGGCCTGGA 1474
|||||
Db 2058 TGGATTATCCCTGCTACTTGAACATGAATTAATGCGCGTGTTCGAAGGTTGA 2117
|||||
QY 1475 AGGAGC 1480
|||||
Db 2118 ATGAGC 2123

Search completed: June 19, 2003, 06:35:34
Job time : 385.433 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 18, 2003, 02:08:24 ; Search time 59.1139 Seconds
(without alignments)
7724.783 Million cell updates/sec

Title: US-09-674-277-1
Perfect score: 1489
Sequence: 1 ctgcagtcggagatgaaag.....ctggaaggagcctgtcgcac 1489

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 150 summaries

Database : Issued Patents-NA:*

- 1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
- 2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
- 3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
- 4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
- 5: /cgn2_6/ptodata/1/ina/PTUS_COMB.seq:*
- 6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	357.6	24.0	2238	2	US-08-674-887A-7
2	357.6	24.0	2238	3	US-08-951-844-7
3	357.6	24.0	2238	4	US-09-412-347-7
4	298.4	20.0	2235	1	US-08-418-782-1
5	298.4	20.0	2235	1	US-08-428-662-1
6	298.4	20.0	2235	2	US-08-852-219-1
7	298.4	20.0	2331	1	US-08-418-782-20
8	298.4	20.0	2331	1	US-08-852-219-20
9	298.4	20.0	4403765	4	US-09-103-840A-2
10	298.4	20.0	4411529	4	US-09-103-840A-1
11	274.2	18.4	2221	1	US-08-418-782-2
12	274.2	18.4	2221	1	US-08-428-662-2
13	274.2	18.4	2221	2	US-08-852-219-2
14	274.2	18.4	4794	2	US-08-459-499-8
15	274.2	18.4	4795	2	US-08-313-185-45
16	274.2	18.4	4795	3	US-09-082-614A-45
17	224.6	15.1	2262	2	US-08-674-887A-5
18	224.6	15.1	2262	3	US-08-951-844-5
19	224.6	15.1	2262	4	US-09-412-347-5
20	48	3.2	7218	1	US-08-232-463-14
21	38.8	2.6	4518	3	US-09-125-287-2
22	38.8	2.6	12839	3	US-09-125-287-1
23	33.8	2.3	579	4	US-08-991-789A-147
24	33.8	2.3	579	4	US-09-062-451-147
25	33.8	2.3	579	4	US-09-598-326-147
26	33.8	2.3	10970	3	US-08-716-351A-5
27	33.4	2.2	1074	4	US-09-134-001C-2054

28	32.4	2.2	6328	4	US-08-913-832A-1	Sequence 1, Appl
29	32.4	2.2	6328	4	US-09-249-181A-1	Sequence 1, Appl
c 30	32.4	2.2	11707	4	US-09-136-574A-1	Sequence 1, Appl
c 31	32.4	2.2	43795	3	US-08-742-185-101	Sequence 101, App
c 32	32.2	2.2	44377	2	US-08-804-227C-7	Sequence 7, Appl
c 33	32.2	2.2	44377	2	US-08-804-198-1	Sequence 1, Appl
c 34	32.2	2.2	2837	4	US-09-136-316-11	Sequence 11, Appl
35	31.8	2.1	2278	4	US-09-442-055-1	Sequence 1, Appl
36	31.8	2.1	3097	4	US-09-282-147-38	Sequence 38, Appl
37	31.6	2.1	2825	4	US-09-196-390-5	Sequence 5, Appl
c 38	31.6	2.1	4353	2	US-08-365-486A-18	Sequence 18, Appl
c 39	31.6	2.1	4353	4	US-08-880-342-18	Sequence 18, Appl
c 40	31.6	2.1	4780	2	US-08-365-486A-20	Sequence 20, Appl
c 41	31.6	2.1	4780	3	US-09-123-708-3	Sequence 3, Appl
c 42	31.6	2.1	4780	3	US-09-123-624-3	Sequence 3, Appl
c 43	31.6	2.1	4780	4	US-08-880-342-20	Sequence 20, Appl
c 44	31.4	2.1	3407	1	US-08-253-155A-7	Sequence 7, Appl
45	31	2.1	289	4	US-09-007-005-17	Sequence 17, Appl
46	31	2.1	289	4	US-09-244-796-17	Sequence 17, Appl
47	30.8	2.1	3153	4	US-09-175-928-9	Sequence 9, Appl
c 48	30.8	2.1	4160	4	US-09-341-678-5	Sequence 5, Appl
49	30.8	2.1	4403765	4	US-09-103-840A-2	Sequence 2, Appl
50	30.8	2.1	4411529	4	US-09-103-840A-1	Sequence 1, Appl
c 51	30.6	2.1	34053	4	US-09-453-702B-96	Sequence 96, Appl
c 52	30.6	2.1	45175	4	US-09-453-702B-116	Sequence 116, App
c 53	30.4	2.0	20199	4	US-08-961-527-6	Sequence 6, Appl
54	30.2	2.0	608	3	US-08-115-753-6	Sequence 6, Appl
c 55	30.2	2.0	849	4	US-09-134-001C-2215	Sequence 2215, App
c 56	30.2	2.0	1316	1	US-08-446-822-7	Sequence 7, Appl
c 57	30.2	2.0	1316	4	US-09-328-314-7	Sequence 7, Appl
c 58	30.2	2.0	1316	5	PCT-US93-12586-7	Sequence 7, Appl
59	30.2	2.0	5398	3	US-09-356-952-11	Sequence 11, Appl
60	30.2	2.0	9848	4	US-09-385-222A-3	Sequence 3, Appl
61	30.2	2.0	13187	4	US-09-422-936-61	Sequence 61, Appl
c 62	30.2	2.0	18627	4	US-08-961-527-113	Sequence 113, App
c 63	30	2.0	500	3	US-09-141-000-2	Sequence 2, Appl
64	29.8	2.0	729	5	PCT-US94-01149-17	Sequence 17, Appl
65	29.8	2.0	1080	4	US-09-615-192A-145	Sequence 145, App
66	29.8	2.0	1229	2	US-08-975-316-85	Sequence 85, Appl
67	29.8	2.0	1229	4	US-09-615-192A-85	Sequence 85, Appl
68	29.8	2.0	4825	6	5459251-1	Patent No. 5459251
69	29.8	2.0	5086	2	US-08-465-485A-19	Sequence 19, Appl
70	29.8	2.0	5086	2	US-08-365-486A-14	Sequence 14, Appl
71	29.8	2.0	5086	3	US-09-080-285-19	Sequence 19, Appl
72	29.8	2.0	5086	4	US-08-880-342-14	Sequence 14, Appl
73	29.8	2.0	5086	4	US-09-724-426-19	Sequence 19, Appl
74	29.8	2.0	5086	4	US-09-233-527-7	Sequence 7, Appl
75	29.8	2.0	5086	5	PCT-US93-05651-4	Sequence 4, Appl
76	29.8	2.0	5086	5	PCT-US93-06251-2	Sequence 2, Appl
77	29.8	2.0	5086	5	US-09-234-186-7	Sequence 7, Appl
78	29.8	2.0	5104	6	5506344-1	Patent No. 5506344
79	29.8	2.0	5449	4	US-09-546-990-1	Sequence 1, Appl
80	29.8	2.0	49377	1	US-08-764-233A-1	Sequence 1, Appl
81	29.6	2.0	1759	4	US-09-491-362-1	Sequence 1, Appl
82	29.6	2.0	1759	4	US-09-874-562-1	Sequence 1, Appl
83	29.6	2.0	3489	2	US-08-728-323A-1	Sequence 1, Appl
c 84	29.6	2.0	3489	4	US-09-298-568-1	Sequence 1, Appl
c 85	29.6	2.0	10348	2	US-08-457-273B-41	Sequence 41, Appl
c 86	29.6	2.0	10348	3	US-08-556-419-13	Sequence 13, Appl
c 87	29.6	2.0	10348	4	US-09-041-886-14	Sequence 14, Appl
c 88	29.6	2.0	10366	1	US-08-246-982A-5	Sequence 5, Appl
c 89	29.6	2.0	10366	1	US-08-453-265-5	Sequence 5, Appl
90	29.6	2.0	32207	2	US-08-770-379-20	Sequence 20, Appl
91	29.6	2.0	32207	4	US-08-757-669A-20	Sequence 20, Appl
92	29.6	2.0	32207	4	US-09-230-371A-20	Sequence 20, Appl
93	29.6	2.0	48908	4	US-09-453-702B-137	Sequence 137, App
c 94	29.4	2.0	409	1	US-08-400-256-37	Sequence 37, Appl
c 95	29.4	2.0	409	3	US-08-975-365-37	Sequence 37, Appl
c 96	29.4	2.0	415	1	US-08-400-256-16	Sequence 16, Appl
c 97	29.4	2.0	415	1	US-08-400-256-22	Sequence 22, Appl
c 98	29.4	2.0	415	1	US-08-400-256-25	Sequence 25, Appl
c 99	29.4	2.0	415	1	US-08-400-256-28	Sequence 28, Appl
c 100	29.4	2.0	415	1	US-08-400-256-31	Sequence 31, Appl

QY 1277 ACGTATGGCCATGATGATGAGGAGTGTGGCCCTGATCGCGGGAGGGGATACATTTGG 1336
Db 768 CCGAATGGCAATGAATGACGAAGAACCGGTGGCTCATAGCGGGTGGACACACCTTCGG 827
QY 1337 TAAAGCACATGGTGCAGCGTCTCTGAAAAATGATTGGGCGAGGCGCTGATGGTGACCC 1396
Db 828 AAAACCCATGGTGTGCGCATGCGGAGAAATATGTGGGCGGAGAGCGCTGCCCGCGAGG 887
QY 1397 TGTGGAGGACAGGAGTGGGATGGAATAAATGTTGTPACAGGAAACGGCAAAATATAC 1456
Db 888 TATTGAAGAAATGAGCTGGGTGGAAAAACACCTACGGCACCGGACACGGTGGCGGATAC 947
QY 1457 CATCACAGTGGCTGGAAGGAGCGTGGTGA 1488
Db 948 CATCACAGTGGACTAGAAGGCGCTGGACCA 979

RESULT 2
US-09-951-844-7
; Sequence 7, Application US/08951844
; Patent No. 6074860
; GENERAL INFORMATION:
; APPLICANT: Robertson et al.
; TITLE OF INVENTION: Catalases
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,
; ADDRESSEE: CECCHI, STEWART & OLSTEIN
; STREET: 6 BECKER FARM ROAD
; CITY: ROSELAND
; STATE: NEW JERSEY
; COUNTRY: USA
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 INCH DISKETTE
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WORD PERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/951,844
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/674,887
; FILING DATE: July 3, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Charles J. Herron
; REGISTRATION NUMBER: 28,019
; REFERENCE/DOCKET NUMBER: 331400-55
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2238 NUCLEOTIDES
; TYPE: NUCLEIC ACID
; STRANDEDNESS: SINGLE
; TOPOLOGY: LINEAR
; MOLECULE TYPE: DNA
US-09-951-844-7

Query Match 24.08; Score 357.6; DB 3; Length 2238;
Best Local Similarity 66.7%; Pred. No. 6.8e-107;
Matches 542; Conservative 0; Mismatches 264; Indels 6; Gaps 2;

QY 683 ATCAATCCCTGGGGGCTGATTTGATTATGTCACAGATTTCAACAGCTGATATGGA 742
Db 168 ATCGACCCAAAGACCGCGGATTTGACTATGCGGAGAGATTTAAGAGCTAGATCTGGC 227
QY 743 GGCTCTGAAAAAGATATCAAGATTTGCTGACAACTTCCAGGATTTGGTGGCCCTGGGA 802
Db 228 AGCGGTAAAAAGACCTGCGCGCTGATGACAGATTCACAGGATTTGGTGGCCAGAGA 287

QY 803 TTATGTCATATTATGTCCTTTCTTTATTCATGATGGCTTGGCAGCGTGGCCGAACATACAG 862
Db 288 TTACGGTCAATTATGTCCTTTCTTTATACCATGATGGCTGGCAGCGGCGGACCTACCG 347
QY 863 GACATATGATGCGCGGGAGGCGCCAGTGTGTGTCAGCAACGTTTCAACCGCTGACACAG 922
Db 348 TATCGGTGATGCGCGGTGGTGGCTCCGGCTCACAGCGCTTCGCGCTCTCAATAG 407
QY 923 CTGCGCGGATTAACCTTAATCTGGATAAAGCCCTCGATTGCTGTGGCCAGTCAAGAAAAA 982
Db 408 CTGCGCAGACAATGCCAATCTGATTAAGCAGCCCTTCTTCTTTGGCCCATCAACAAAA 467
QY 983 ATACGGCTCCAGTATTTCCTGGGAGAGACCTGATGGTCTGACCTGATGTAATGTCCTTGA 1042
Db 468 ATACGGTTCGAAAAATCTCTGGGCGGATCTAATGATACCTACAGGAAACCTAGCTCTGGA 527
QY 1043 ATCCATGGGATTTAAAAACGCTGGGATTGCTGGCGGAAGAGATGACCTGGGAGTTC --- 1099
Db 528 AACTATGGGCTTTAAACCTTTTGGTTTGGAGTGGCAGAGCAGATGTATGGAGCGCTGA 587
QY 1100 GGACCTGGTATCTGCGGGCTTGACAAACAGCCCTTTGCGAGATAAACCGGATATAA --- AA 1156
Db 588 AGAAGATGATATCTGCGGAGCAGAAACCGAATGGCTGGGAGACAGCGCTATGAAGTGA 647
QY 1157 CGGGAACCTCAGAAACCTTTGCGCGCAGCAGAGATGGGACTATTATGTCATGCTGA 1216
Db 648 CCGAGAGCTCGAAAAATCCCTGGGAGCGGTACAAATGGGACTCATCTATGTAACCCGGA 707
QY 1217 AGGCGCGGTGGAAAAACCATCTCTGCGCTTCGCGGAAAGATATCAGGGAAGCTTTTC 1276
Db 708 AGGACCCACGCGCAGCAGCCCTATCGCTGCTGCGGTGATTCGCTGAGACTTTTGG 767
QY 1277 ACGTATGGCATGATGATGAGGAGACTGTGGCCCTGATCGCGGAGGGGATACATTGG 1336
Db 768 CCGAATGGCAATGAATGACGAAGAACCGGTGGCTCATAGCGGTGGACACACCTTCGG 827
QY 1337 TAAAGCACATGTCAGCGCTCTCTGAAAAATGTTGCGGAGGCGCTGATGTCGACCC 1396
Db 828 AAAAACCCATGGTGTGCTCGCGATCGGAGAAATATGTGGCGGAGAGCGCTGCCCGCGAGG 887
QY 1397 TGTGGAGGAGCAGGAGTGGGATGGAAAAATAATGTTGTPACAGGAAACGGCAAAATATAC 1456
Db 888 TATTGAAGAAATGAGCTGGGTGGGAGAAAAACACCTACGGCCCGGACACGGTGGGATAC 947
QY 1457 CATCACAGTGGCTGGAAGGAGCGCTGGTGA 1488
Db 948 CATCACAGTGGACTAGAAGGCGCTGGACCA 979

RESULT 3
US-09-412-347-7
; Sequence 7, Application US/09412347
; Patent No. 6410290
; GENERAL INFORMATION:
; APPLICANT: Robertson, Dan E.
; APPLICANT: Sanyal, Indrajit
; APPLICANT: Adhikari, Robert S.
; TITLE OF INVENTION: CATALASES
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: US
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/412,347

```

; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/674,887
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Haile, Ph.D., Lisa A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 09015/002001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619/678-5070
; TELEFAX: 619/678-5099
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2238 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 1...2235
; US-09-412-347-7

Query Match      24.0%; Score 357.6; DB 4; Length 2238;
Best Local Similarity 66.7%; Pred. No. 6.8e-107;
Matches 542; Conservative 0; Mismatches 264; Indels 6; Gaps 2;

QY 683 ATCAATCTCTGGGGGCTGATTTGATATGCCACAGATTTCACACGCTGGATATGGA 742
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 168 ATCGGACCCAAAGACCGCGGATTTGACTATCCGAGAGTTTAAAGAGCTAGATCTGC 227
QY 743 GGTCTCGAAAAGATATCAAGATTGCTGACACTTCCAGACTTGGTGCCTCGCGGA 802
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 228 AGCGGTAAAGAGACCTGGACGCTTAATGACGATTCAGAGCTGGTGGCCAGCAGA 287
QY 803 TTATGCTCATATGTCCTTTCTTTATTCGATGCTGGCAGCTGCGGAAACATACAG 862
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 288 TTACGGTCTATTATGCTCTCTTTATACGATGCGTGGCAGCGCCGCACTACCG 347
QY 863 GACATATGATGCCGGGAGCGGCGAGTGGTGGTGCAGCAACGTTTGAACGCTGAACAG 922
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 348 TATCGGTGATGCCGTGGTGGCGTGGCTCCGCTCAGCGCTTCGCGCTCTCAATAG 407
QY 923 CTGGCCGAGTAACTTAATCTGGATAAAGCCGCTCGATTGCTGGCCAGTCAAGAAAA 982
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 408 CTGGCCAGACATGCAATCTGGATAAAGACGCTTGTCTTTTGGCCATCAACAAAA 467
QY 983 ATACGCTCCAGTATTTCTCTGGGAGACCTGATGGTCTGACTGTTATGTTGCCCTGA 1042
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 468 ATACGGTCAAAAATCTCTGGGCGGATCTAATGATCTCAGAGAAACGATGCTCTGA 527
QY 1043 ATCCATGGGATTTAAACGCTGGGATTTGCTGGCGGAGAGAGATGACTGGAGTC--- 1099
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 528 AACTATGGCTTTAAACTTTTGGTTTTCAGGTGGCAGAGATGATATGGAGCCTGA 587
QY 1100 GGACTGTGATACCTGGGGGCTGACAAAGCCTCTTTCAGATAACCGGATATAA---AA 1156
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 588 AGAAGATGATACCTGGGAGCAGAAACCGAATGCTGGGAGACAGCGCTATGAAGTGA 647
QY 1157 CGGNAATCTCAGAAACCTCTTCCGCCAGCAGAGATGGGACTTATTATCTCAATCTGA 1216
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 648 CCGAGAGTCAAAATCCCTGGGAGCGGTACAAATGGGACTCATCTATGTAACCCCGA 707
QY 1217 AGGCCCCGGTGAAGAACCATCTCTGCTTCCGCAAGAGATATCAGGAACTTTTTC 1276
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 708 AGGACCCAGCGGACGACACCTATCGCTGCTGGCGGTGATATTCGTGAGACTTTTG 767
QY 1277 ACGTATGGCATGATGATGAGGACTGTGGCCCTGATCCGCGGAGGCGCATACATTTGG 1336
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 768 CCGAATGGCAATGAATGACGAAGAAACCGTGGCTCTCATAGCGGTGGACACACTTCG 827
QY 1337 TAAGACACATGTTGACGCTCTCTGAAATAATGTTGGCGAGGCGCTGATGGTCAAC 1396
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

DB 828 AAAAAACCATGGTCTGCCGATGCGGAGAAATATGTGGCGGAGAGCCTGCCGCCGAGG 887
QY 1397 TGTGAGGAGGAGGAGCTGGGATGATGATGATGATGATGATGATGATGATGATGATGAT 1456
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 888 TATTGAGAAATGAGCTTGGGTGGGAGAAACACCTACGCGACCGACACGCTGGCGATAC 947
QY 1457 CATCACCAGTGGCTTGGGAGGAGGCTGTGCTGA 1488
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 948 CATCACCAGTGGACTAGAGAGCGCTGGACCA 979

RESULT 4
US-08-418-782-1
; Sequence 1, Application US/08418782
; Patent No. 5658733
; GENERAL INFORMATION:
; APPLICANT: Cockerill, Franklin R.
; APPLICANT: Kline, Bruce C.
; APPLICANT: Uhl, James R.
; TITLE OF INVENTION: Detection of Isoniazid Resistant Strains
; TITLE OF INVENTION: of M. Tuberculosis
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Schwegman, Lundberg & Woessner
; STREET: 3500 IDS Center
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/418,782
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Woessner, Warren D.
; REGISTRATION NUMBER: 30,440
; REFERENCE/DOCKET NUMBER: 150.141US1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-339-0331
; TELEFAX: 612-339-0361
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2235 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; US-08-418-782-1

Query Match      20.0%; Score 298.4; DB 1; Length 2235;
Best Local Similarity 62.4%; Pred. No. 1.8e-87;
Matches 502; Conservative 0; Mismatches 296; Indels 6; Gaps 2;

QY 688 ATCCCTGGGGGCTGATTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 747
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 176 ACCCGATGGTGGGGGTTTCGACTATGCGCGGAGGTCGCGACCATGACGCTTGACGCC 235
QY 748 TGAAGAAAGATATCAAGATTTGCTGACAACTTCCAGGATTTCCAGGATTTGCTGGGATG 807
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 236 TGACGCGGACATCGAGGAAGTATGACCACTCCGACCGTGGTGGCCCGCGGACTAG 295
QY 808 GTCATATGGTCTTCTTTTATCTGATGCTTGGCTGGCAGGTCGCGGAACATACAGACAT 867
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 296 GCCACTACGGCGCTGTTTATCCGATGGCTGGCAGCTGCCGCACTACCGCATCC 355
QY 868 ATGATGGCGGGGAGCGCGCTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 927
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 356 ACGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 415

```


RESULT 6

US-08-852-219-1
; Sequence 1, Application US/08852219
; Patent No. 5922575
; GENERAL INFORMATION:
; APPLICANT: Cockerill, Franklin R.
; APPLICANT: Kline, Bruce C.
; APPLICANT: Uhl, James R.
; TITLE OF INVENTION: Detection of Isoniazid Resistant Strains
; TITLE OF INVENTION: of M. Tuberculosis
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Muehling, Raasch & Gebhardt, P.A.
; STREET: 119 No. 5922575th Fourth Street, Ste. 203
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/852,219
; FILING DATE: 07-May-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Sandberg, Victoria A.
; REGISTRATION NUMBER: 41,287
; REFERENCE/DOCKET NUMBER: 230.00010130
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-305-1226
; TELEFAX: 612-305-1228
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2235 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
US-08-852-219-1

Query Match 20.0%; Score 298.4; DB 2; Length 2235;
Best Local Similarity 62.4%; Pred. No. 1.8e-87;
Matches 502; Conservative 0; Mismatches 296; Indels 6; Gaps 2;

QY	688	ATCCCTGGGGGCTGATTTTGAATATCCACACAGATTTCACACAGCTGGATATGGAGCTC	747
DB	176	ACCGGATGGTGGCGCGCTTCGACTATCGCGGAGTCCGACCATCGAGCTTGACGCC	235
QY	748	TGAAAAAGATATCAAGATTTCGACAACTTCCAGGATTTGGTCCCTGGCGATATG	807
DB	236	TGACGGGGACATCGAGAAAGTATGACACCTCGCAGCGGTGGTGGCCGCGACTAG	295
QY	808	GTCATTATGTCCTTTCTTTATTCGATGCTTGGCACGCTGCGGAACATACAGACAT	867
DB	296	GCCACTAGGGCGCTGTTATCGGATGGCTGGCACGCTGCGGACCTACCGATCC	355
QY	868	ATGATGCGGGGAGGCGCCAGTGGTGGTCAGAACGTTTTGAACGCTGAACAGCTGC	927
DB	356	ACGACGCGCGGGCGCGCGGGGCGCATCGAGCGGTTCGGCGCTTAACAGCTGC	415
QY	928	CGGATAACGTTAATCTGGAATAAGCCGCTGATGCTGTGGCCAGTCAGAAAAATACG	987
DB	416	CCGACACCGCCAGTTGGACAAGCGCGCGGCTGCTGTGGCGGTCACAGAAATACG	475
QY	988	GCTCCAGTATTCCTGGGGAGACCTGATGCTCCTGACTGATATGTTGCCCTTGAATCA	1047
DB	476	GCAAGAGCTCTCATGGCGGACCTGATGTTTCCCGGACACTGCGGCTGGATCA	535
QY	1048	TGGGATTTAAACGCTGGGATTTGCTGGCGGAAGAAAGATGACTGGGAGTCGGACCTG	1107

DB	536	TGGCTTCAAGACGCTTCGGGTTTCGGCTTCGGCGGGTCCGACAGTGGGAGCCCGATGAGG	595
QY	1108	TATACTGGGGCCTTGACACAAGCCTTTTCAGATACACCGGATA---AAAACGGGAAC	1164
DB	596	TCTATTGGGCAAGAAAGCCACCTGGCTCGCGATGAGCTTACAGCGGTAAAGCGGATC	655
QY	1165	TTCAAGAAACCTTTTCGGCCACGACAGATGGAGCTTATTATGTCATCTCTGAAGGCCCG	1224
DB	656	TGGAAACCGCTGGCGCGGTGAGATGGGGTGATCTACGTGAACCCGAGGGGCCGA	715
QY	1225	GTGAAACACGATCTCTCTGGCTTCGCGAAAGATATCAGGAAGCTTTTTCACGTATGG	1284
DB	716	ACGCAACCGGACCCCATGCGCGGTCCACATTCGCGAGAGTTTCGCGGATGG	775
QY	1285	CCATGATGATGAGGAGACTGTGCCCTGATCGCGGAGGACATACATTGTGTAAGCAC	1344
DB	776	CCATGAACGAGCTCGAAACAGCGCGCTGATCTCGCGGTACACTTTCGGTAAGACCC	835
QY	1345	ATGCTGCAGCGTCTCTCTGAAAAATGATTGGCGCAGGCGCTGATGTCACCTGTGGAGG	1404
DB	836	ATGCGCGG---GCCCGCGGATCTGTGCGGCCCGAACCCGAGGCTGCTCGCTGGAGC	892
QY	1405	AGCAGGACTGGATGGAATAAATGTTGTTACAGGAACCGCAATATACCATCACCA	1464
DB	893	AGATGGGCTTGGCTGGAAGAGCTCGTATGGCACCGAACCGGTAAAGGACCGATCACCA	952
QY	1465	GTGCGCTGGAAGGAGCTGTGCGA	1488
DB	953	GCGCATCGAGTGTGATGACGA	976

RESULT 7

US-08-418-782-20
; Sequence 20, Application US/08418782
; Patent No. 5658733
; GENERAL INFORMATION:
; APPLICANT: Cockerill, Franklin R.
; APPLICANT: Kline, Bruce C.
; APPLICANT: Uhl, James R.
; TITLE OF INVENTION: Detection of Isoniazid Resistant Strains
; TITLE OF INVENTION: of M. Tuberculosis
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Schwegman, Lundberg & Woessner
; STREET: 3500 IDS Center
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE:
; APPLICATION NUMBER: US/08/418,782
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Woessner, Warren D.
; REGISTRATION NUMBER: 30,440
; REFERENCE/DOCKET NUMBER: 150.141US1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-339-0331
; TELEFAX: 612-339-3061
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2331 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; FEATURE:

NAME/KEY: CDS
LOCATION: 70...2289
US-08-418-782-20

Query Match 20.0%; Score 298.4; DB 1; Length 2331;
Best Local Similarity 62.4%; Pred. No. 1.9e-87;
Matches 502; Conservative 0; Mismatches 296; Indels 6; Gaps 2;

```
QY 588 ATCCCTGGGGGGCTGATTTTGATATGCCACACAGATTTCAACAGCTGGGATATGAGGCTC 747
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 236 ACCCGATGGGTGGGGGCTTCGACTATGCCGCGAGGTCCGACCATCGACGTTGACGCC 295
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 748 TGAATAAAGATATCAAAAGATTGCTGACAACTTCCACAGATTGTCGCCCTGCGGATTATG 807
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 296 TGACGGGACATCGAGGAAGTATGACCACTTCGACCGTGGTGCCGCCGCGACTACG 355
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 808 GTCAATTATGTCCTTTCTTTATTGATGCTTGGCAGCTGGCGGAACATACAGGACAT 867
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 356 GCCACTACGGGCGGCTGTTTATCCGATGGCTGGCAGCTGCCGACCTACCGCATCC 415
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 868 ATGATGCCGGGAGGCGGCTGGTGTGACCAAGCTTTTGAACCGCTGAACAGCTGGC 927
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 416 ACACGCGCGCGCGCGCGGGGGCGGATGCGAGCGGTTTCGCGCCCTTAAACAGCTGCG 475
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 928 CGGATACGTTAATCTCGATAAAGCCGCTGATTCCTCCAGAGTTCGCGGAGTTCGCGGATCC 987
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 476 CCACACGCGGACCTTGACACAGCGCGCGCGGCTGCTGTGGCCGCTCAAGAGATCG 535
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 988 GCTCCAGATATTCTCGGGGAGACGTATGATGTCCTGACTGTGTAATGTTGCCCTTGAATCCA 1047
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 536 GCAAGAGCTCTCATGGCGGACTGATTGTTTCGCGGCACTGCGCGCTGGAATCGA 595
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1048 TGGGATTTAAACGCTGGGATTTGCTGGCGGAGAGAGATGACTGGGAGTCGACCTGG 1107
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 596 TGGGCTTCAAGAGCTTCGGGTTGCGGTTTCGCGCGGCTGACAGCTGGGAGCGGATGAG 655
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1108 TATATCGGGGCGCTGACAAAGCCCTCTTCAGATACCGGATA---AAACGGGAAC 1164
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 556 TCTATTGGGCAAGGAAGCCACCTTCGGCTCGCGGATGAGCTTACAGCGGTGAGCGGATC 715
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1165 TTCAGAAACCTCTTGGCGGCGCACAGATGGGACTTATTTATGCAATTCCTGAAGGCCCG 1224
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 716 TGGAGAACCGCTTGGCGGCGGTGCAGATGGGCTGATCTACGTGAACCCGAGGGGCGGA 775
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1225 GTGGAACACAGATCTCTGCTTCCGCGAAGATATCAGGGAAGCTTTTACAGTATGG 1284
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 776 ACGGCAACCGGACCCCATGGCGCGCGGTGACATTCGCGAGACGTTTCGCGCATGG 835
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1285 CCATGGATGATGAGGAGCTGTGGCCTGATCGCGGAGGCGCATACATTTGTTAAAGCAC 1344
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 836 CCATGAACGACGTGGAACAGCGCGCTGATGTCGCGGTTCACACTTTCGTTAAGACC 895
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1345 ATGCTGACGCTCTCTTGAATAATGTTGGCGAGGCGCTGATGTGCACCTGTGGAGG 1404
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 896 ATGGCGCG---GCCCGCGCATCTGTGTGGCGCGGACCGAGGCTGCTCCGCTGGAGC 952
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1405 AGCAGGACGTGGATGGAATAATGTTGTCAGGAACGCGCAATATACCATCA 1464
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 953 AGATGGGCTTGGCTGGAGAGCTGCTGATGGCAGCGGACCGGTAAGGACGCGATCA 1012
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1465 GTGGCTTGAAGGAGCTGTGTCGA 1488
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1013 CGGCGATCGAGGTCGTATGACGA 1036
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

RESULT 8

US-08-852-219-20
; Sequence 20, Application US/08852219
; Patent No. 5922575
; GENERAL INFORMATION:
; APPLICANT: Cockerill, Franklin R.
; APPLICANT: Kline, Bruce C.
; APPLICANT: Uhl, James R.

TITLE OF INVENTION: Detection of Isoniazid Resistant Strains
TITLE OF INVENTION: of M. Tuberculosis
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: Muetting, Raasch & Gebhardt, P.A.
STREET: 119 No. 5922575th Fourth Street, Ste. 203
CITY: Minneapolis
STATE: MN
COUNTRY: USA
ZIP: 55401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/852,219
FILING DATE: 07-May-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Sandberg, Victoria A.
REGISTRATION NUMBER: 41,287
REFERENCE/DOCKET NUMBER: 230.00010130
TELECOMMUNICATION INFORMATION:
TELEPHONE: 612-305-1226
TELEFAX: 612-305-1228
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 2331 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
FEATURE:
NAME/KEY: CDS
LOCATION: 70...2289
US-08-852-219-20

Query Match 20.0%; Score 298.4; DB 2; Length 2331;
Best Local Similarity 62.4%; Pred. No. 1.9e-87;
Matches 502; Conservative 0; Mismatches 296; Indels 6; Gaps 2;

```
QY 588 ATCCCTGGGGGGCTGATTTTGATATGCCACACAGATTTCAACAGCTGGGATATGAGGCTC 747
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 236 ACCCGATGGGTGGGGGCTTCGACTATGCCGCGAGGTCCGACCATCGACGTTGACGCC 295
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 748 TGAATAAAGATATCAAAAGATTGCTGACAACTTCCACAGATTGTCGCCCTGCGGATTATG 807
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 296 TGACGGGACATCGAGGAAGTATGACCACTTCGACCGTGGTGCCGCCGCGACTACG 355
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 808 GTCAATTATGTCCTTTCTTTATTGATGCTTGGCAGCTGGCGGAACATACAGGACAT 867
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 356 GCCACTACGGGCGGCTGTTTATCCGATGGCTGGCAGCTGCCGACCTACCGCATCC 415
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 868 ATGATGCCGGGAGGCGGCTGGTGTGACCAAGCTTTTGAACCGCTGAACAGCTGGC 927
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 416 ACACGCGCGCGCGCGCGGGGGCGGATGCGAGCGGTTTCGCGCCCTTAAACAGCTGCG 475
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 928 CGGATACGTTAATCTCGATAAAGCCGCTGATTCCTCCAGAGTTCGCGGAGTTCGCGGATCC 987
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 476 CCACACGCGGACCTTGACACAGCGCGCGGCTGCTGTGGCCGCTCAAGAGATCG 535
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 988 GCTCCAGATATTCTCGGGGAGACGTATGATGTCCTGACTGTGTAATGTTGCCCTTGAATCCA 1047
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 536 GCAAGAGCTCTCATGGCGGACTGATTGTTTCGCGGCACTGCGCGCTGGAATCGA 595
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1048 TGGGATTTAAACGCTGGGATTTGCTGGCGGAGAGAGATGACTGGGAGTCGACCTGG 1107
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 596 TGGGCTTCAAGAGCTTCGGGTTGCGGTTTCGCGCGGCTGACAGCTGGGAGCGGATGAG 655
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1108 TATATCGGGGCGCTGACAAAGCCCTCTTCAGATACCGGATA---AAACGGGAAC 1164
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 556 TCTATTGGGCAAGGAAGCCACCTTCGGCTCGCGGATGAGCTTACAGCGGTGAGCGGATC 715
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```

QY 1165 TTCAGAACCTCTTTCGCGCACGACAGATGGGACTTATTTATGTCATCTCTGAAGGCCCG 1224
Db 716 TGGAGAACCCGCTGGCGCGGTGCAGATGGGCTGATCTACGTGAACCCGGAGGGCCGA 775
QY 1225 GTGGAACACAGATCTCTGGCTTCGCGGAAGATATCAGGAAGCTTTTTCACGTATGG 1284
Db 776 ACGCAACCCGGACCCCATGAGCGCGGTGCACATTCGCGAGACGTTTCGCGCGATGG 835
QY 1285 CCATGATGATGAGGAGACTGTGGCCCTGATCGCGGAGGCGATACATTTGCTAAAGCAC 1344
Db 836 CCATGAGACGTCGAACACAGCGCGCTGATGTCGCGCGTCCACATTTTCGCTAAGACCC 895
QY 1345 ATGGTCGAGGCTCTCTGAAAATGATTGGCGCAGGCGCTGATGTCGACCTGTGGAGG 1404
Db 896 ATGGCGCGG---GCCCGCGCGATCTGTCGCGCGCGAACCAGCGCTGCTCGCGTGAGC 952
QY 1405 AGCAGGAGACTGGGATGGAATAAATGTTGACAGGAACGCGAATATACCATCACC 1464
Db 953 AGATGGCTTGGCTGGAAGAGCTGTATGGCACCAGGAAACCGGTAAGGACGCGATCACC 1012
QY 1465 GTGGCTGGAAGAGCGCTGTGCGA 1488
Db 1013 GCGGCATCGAGTCTGATGGACGA 1036

RESULT 9
US-09-103-840A-2/c
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

Query Match 20.0%; Score 298.4; DB 4; Length 4403765;
Best Local Similarity 62.4%; Pred. No. 1.8e-85;
Matches 502; Conservative 0; Mismatches 296; Indels 6; Gaps 2;

QY 688 ATCCCTGGGGCGCTGATTTGATTCATCCACAGATTTCAACAGCTGGATGAGGCTC 747
Db 2153203 ACCCGATGGTGGCGGTTCGACTATGCCGGAGGTGCGACCATCGACTTGACGCC 2153144
QY 748 TGAAGAAGATATCAAGATTGCTGACACTTCCAGGATTTGGTCCCTCGCGGATATG 807
Db 2153143 TGACCGGGACATCAGGAAGTGTATGACCACCTCGCAGCGGTGGTGGCGCGGACTACG 2153084
QY 808 GTCATATGTCCTTTCTTATTCGTATGCTTGGACGCTGGCGGACATACAGACAT 867
Db 2153083 GCCACTAGGCGCGCTGTTATCGGATGGGTGGCAGCTGCGCGACCTACCGCATCC 2153024
QY 868 ATGATGCGCGGGAGCGCCAGTGGTGTGAGCAACCTTTTGAACCGCTGAACAGCTGC 927
Db 2153023 ACGACGCGCGCGCGCGCGGCGGATGACGCGGTTCGCGCGCTTAACAGCTGCG 2152964
QY 928 CGGATACGTTATCTGGATAAAGCCGCTGATGCTGTGGCCAGCTCAAGAAAAAATACG 987

```

```

Db 2152963 CGACACGCCAGCTTGGACAGGCGCGCGCTGCTGTGCGGCTCAAGAAGATACG 2152904
QY 988 GCTCCAGTATTTCTTGGGAGACCTGATGTCCTGACTTGGTAAATGTTGCCCTTCAATCCA 1047
Db 2152903 GCAAGAAGCTCTCATGGGCGACCTGATGTTTTCGCGGCACTGCGCGCTGGAATCGA 2152844
QY 1048 TGGATTTAAACCGCTGGGATTTCTGCGCGAAGAGATGATCTGGAGTTCGAGCTGG 1107
Db 2152843 TGGGCTTCAAGACCTTCGGGTTCCGGCTTCGCGCGGCTCGACCACTGGGAGCGGATGAGG 2152784
QY 1108 TATACTGGGGCGCTGACAAAGCCTCTTCAGATACCGGATA---AAAACGGGAAC 1164
Db 2152783 TCTATTGGGCAAGAAAGCCACCTGGCTCGCGATGAGCGTTACAGCGGTAAGCGGATC 2152724
QY 1165 TTCAGAAACCTTTTCGCGCACGAGATGGGACTTATTTATGTCATCTTGAAGGCCCG 1224
Db 2152723 TGGAGAACCCGCTGGCGCGGTGCAGATGGGCTGATCTACGTGAACCCGGAGGGCCGA 2152664
QY 1225 GTGGAACACAGATCTCTGCTTCCGCGAAGATATCAGGAAGCTTTTTCACGTATGG 1284
Db 2152663 ACGCAACCCGCGACCCCATGCGCGCGCTGCACATTCGCGAGACGTTTCGCGCGATGG 2152604
QY 1285 CCATGATGATGAGGAGACTGTGGCCCTGATCGCGGAGGCGCATACATTTGTTAAAGCAC 1344
Db 2152603 CCATGAACGAGCTCGAAACAGCGCGCTGATCTCGCGGTGCACATTTTCGGTAAGACC 2152544
QY 1345 ATGGTCGAGCTCTCTGAAAATGATTGGCGCAGGCGCTGATGTCGACCTGTGGAGG 1404
Db 2152543 ATGGCGCG---GCCCGCGCGATCTGGTGGCGCGCGAACCAGGCTGCTCCGCTGGAGC 2152487
QY 1405 AGCAGGAGCTGGGATGGAATAAATGTTGTCAGGAACGCGCAATATACCATCACC 1464
Db 2152486 AGATGGCTTGGCTGGAAGAGCTCTGATGGCACCGGACCGGTAAGGACGCGATCACC 2152427
QY 1465 GTGGCTGGAAGAGCGCTGTGCGA 1488
Db 2152426 GCGGCATCGAGTCTGATGGACGA 2152403

RESULT 10
US-09-103-840A-1/c
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv
US-09-103-840A-1

Query Match 20.0%; Score 298.4; DB 4; Length 4411529;
Best Local Similarity 62.4%; Pred. No. 1.8e-85;
Matches 502; Conservative 0; Mismatches 296; Indels 6; Gaps 2;

QY 688 ATCCCTGGGGCGCTGATTTGATTCATCCACAGATTTCAACAGCTGGATGAGGCTC 747
Db 2155943 ACCCGATGGTGGCGGTTCGACTATGCCGGAGGTGCGCGGAGGTGCGACCTGACGCTG 2155884
QY 748 TGAAGAAAGATATCAAGATTTGTCGACAACTTCCAGGATTTGGTCCCTCGCGATATG 807

```


QY 1406 GCAGGACTGGATGGAATAAATGTTGTACAGGAACGGCAAAATATACCATCACAG 1465
 Db 891 GATGGCTTGGCTGGGAAGAGCTCGTATGGCACCGGAACGGTAAGACCGCATCACAG 950
 QY 1466 TGGCTGGAAGGAGCTGGTCGA 1488
 Db 951 CGGATCGAGTGTATGGACGA 973

RESULT 12
 US-08-228-662-2
 ; Sequence 2, Application US/08228662
 ; Patent No. 5688639
 ; GENERAL INFORMATION:
 ; APPLICANT: COCKERILL, FRANKLIN R.
 ; APPLICANT: KLINE, BRUCE C.
 ; APPLICANT: UHL, JAMES R.
 ; TITLE OF INVENTION: DETECTION OF ISONIAZID RESISTANT STRAINS
 ; TITLE OF INVENTION: OF M. TUBERCULOSIS
 ; NUMBER OF SEQUENCES: 7
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: SCHWEGMAN, LUNDBERG & WOESSNER, P.A.
 ; STREET: 3500 IDS CENTER
 ; CITY: MINNEAPOLIS
 ; STATE: MINNESOTA
 ; COUNTRY: USA
 ; ZIP: 55402
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/228,662
 ; FILING DATE: 18-APR-1994
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: WOESSNER, WARREN D.
 ; REGISTRATION NUMBER: 30,440
 ; REFERENCE/DOCKET NUMBER: 150.123US01
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 612-339-0331
 ; TELEFAX: 612-339-3061
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 221 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: DNA (genomic)
 ; US-08-228-662-2

Query Match 18.4%; Score 274.2; DB 1; Length 2221;
 Best Local Similarity 61.4%; Pred. No. 1.6e-79;
 Matches 493; Conservative 0; Mismatches 303; Indels 7; Gaps 3;

QY 688 ATCCCTGGGGGCTGATTTGATTATGCCACAGATTTCACAGCTGGATGGAGCTC 747
 Db 176 ACCCGATGGTGGCGTTCGACTATGCCGGGAGTCCGACAGTCGACTTGAAGCC 235
 QY 748 TGAAGAAGATATCAAGATTGCTGACAACTTCCAGGATTGTGCCCTCGCGGATTATG 807
 Db 236 TCACGCGGACATCGAGAGTGTATGACCACTCGCAGCGTGTGGCGCGCGACTACG 295
 QY 808 GTCAATTATGTCCTTTCTTTATTCGATGGCTGGCAGCGTGGCGGAACATACAGGACAT 867
 Db 296 GCCACTAGGCGCGCTGTTATCCGAGTGGGTGGCGACGCTGGCGCACCTACCGCATCC 355
 QY 868 ATGATGCGCGGGAGCGCCAGTGGTGTGAGCAACGTTTGAACCGCTGAACAGCTGCC 927
 Db 356 ACGACGCGCGGGCGCGCGCGGCGGCGCATGACGCGTTCGCGCGCTTAACAGCTGGC 415
 QY 928 CGGATACGTTAATCTGGATAAAGCCGCTCGATTGCTGTGGCGAGTCAAGAAAAATACG 987

Db 416 CGACACGCCAGCTTGGCAAGCGCGCGCTGCTGTGGCCGCTCAAGAAGAAGTACG 475
 QY 988 GCTCCAGTATTTCTTGGGGAGACCTGATGGTCTGACTGGTAA--TCTTGCCTTTGAATC 1045
 Db 476 GCAAGAGCTCTCATGGCGGACCTGATTTTCGCGCGCAACCGCTGCGTGGGAATC 535
 QY 1046 CATGGGATTTAAACGCTGGGATTGCTGGCGGAAGAGAAAGATGACTGGAGTGGACCT 1105
 Db 536 GATGGGCTTCAAGACGTTTCGGGTTTCGGCTTCGG--CGTTCGACAGTGGGAGACCATGA 593
 QY 1106 GGTATCTTGGGGGCTGACACAAAGCCTCTTGCAGATAACCGGATATAAACCGGAACCT 1165
 Db 594 GGTCTATTGGGCAAGAACCCCTCGCTCGCGGATGACGGTTACAGCGTAAGCGATCT 653
 QY 1166 TCAGAAACCTCTTCCGCCACGACAGATGGGACTTATTTATGTCAATCTTGAAGCCCGG 1225
 Db 654 GGAGAACCCGCTGGCGCGGTGCAGATGGGCTGATCTAGTGNACCCGAGGCGCGAA 713
 QY 1226 TGGAAACACAGATCTCTGCTTCCGCGAAGATATCAGGGAAGCTTTTTCAGTATGC 1285
 Db 714 CGCAACCCGACCCCATGGCGCGGTGCGACATTCGCGAGACGTTTCGGCGCATGGC 773
 QY 1286 CATGGATGATGAGAGACTGTGCCCTGTATCGCGGAGGCGCATACATTTGTAAACACA 1345
 Db 774 CATGAACGACGTGGAACACGCGCTGATCGTGGCGGTACACACTTTCGGTAAGACCA 833
 QY 1346 TGGTGCAGCTCTCTGTAATAATGTAATGGCGCAGGCGCTGATGGTGCACCTTGGAGGA 1405
 Db 834 TGGCGCG---GCCCGCGCATCTGTGCGCCCGCAACCGAGGCTGCTCCGCTGGAGCA 890
 QY 1406 GCAGGAGCTGGGATGGAATAAATAATGTGTGACAGGAACGGCAAAATATACCATCACAG 1465
 Db 891 GATGGGCTTGGCTTGAAGAGCTCGTATGCCACCGGACCGGTAAGACCGCATCACAG 950
 QY 1466 TGGCTGGAAGGAGCTGTGTCGA 1488
 Db 951 CGGATCGAGTGTATGGACGA 973

RESULT 13
 US-08-252-219-2
 ; Sequence 2, Application US/08852219
 ; Patent No. 5922575
 ; GENERAL INFORMATION:
 ; APPLICANT: COCKERILL, Franklin R.
 ; APPLICANT: Kline, Bruce C.
 ; APPLICANT: Uhl, James R.
 ; TITLE OF INVENTION: Detection of Isoniazid Resistant Strains
 ; TITLE OF INVENTION: of M. Tuberculosis
 ; NUMBER OF SEQUENCES: 22
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Muetting, Raasch & Gebhardt, P.A.
 ; STREET: 119 No. 5922575th Fourth Street, Ste. 203
 ; CITY: Minneapolis
 ; STATE: MN
 ; COUNTRY: USA
 ; ZIP: 55401
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/852,219
 ; FILING DATE: 07-May-1997
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Sandberg, Victoria A.
 ; REGISTRATION NUMBER: 41,287
 ; REFERENCE/DOCKET NUMBER: 230.00010130
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 612-305-1226

Query Match	18.4%	Score 274.2	DB 2	Length 2221
Best local Similarity	61.4%	Pred. No. 1.6e-79		
Matches 493	Conservative 0	Mismatches 303	Indels 7	Gaps 3
Qy	688	ATCCCTGGGGCGCTGATTTTGGATTATATGCCACCAAGATTTCAACAGCTGGATATATGAGGCTC	747	
Db	176	ACCCGATGGGTGCGGGCTTCGACTATGCCGCGAGGTTCGGACCACTGCAGCTTGAAGGCC	235	
Qy	748	TGAAAAAGATATCAAGATTTCTGTGACAACTTTCCACAGGATTTGGTGCCTTGCAGATATG	807	
Db	236	TGACGCGGGACATCGAGGAAGTATGATACCACTTCGACGCGCTGGTGCCTCCGCGCACTACG	295	
Qy	808	GTCAATTATGTCTCTTTTATTCGTATGGCTTGGCACGCTGCCGGAACATACAGACAT	867	
Db	296	GCCACTACGGCCGCTGTTTATCCGATGCGCTGGCACGCTGCCGCACTCCGCATCC	355	
Qy	868	ATGATGCCGGGAGGCGCCAGTGGTGGTCAGCAACCTTTTGAACCGCTGAACAGCTGGC	927	
Db	356	ACGACGCCGCGCGCGCGGGCGGCATGCAGCGTTTCGCGCGCTTAACAGCTGGC	415	
Qy	928	CGGATAACGTTAATCTGGATTAAGCCCGTCGATTTGCTGTGGCCACGTCAAGAAAAATACG	987	
Db	416	CCGACAAACGCCACTTGGACAAAGCGCGCGGCTGCTGTGGCCGGTCAAGAAAGATACG	475	
Qy	988	GCCTCAGTATTTCTTGGGAGACCTGATGTCTCTGACTGGTAA--TGTTGCCCTTGAATC	1045	
Db	476	GCAAGAAGCTCTCATGGCGGACCTGATTTGTTTCCGCGCAACCGTCGGCTCGGAATC	535	
Qy	1046	CATGGGATTTAAACGCTGGGATTTGCTGCGCGAAGAAAGATGACTGGAGTCGGACCT	1105	
Db	536	GATGGGCTTCAGACGCTCGGGTTCGGCTTCGG--GCGTCGACCACTGGGAGACCGATGA	593	
Qy	1106	GGTATACTGGGGGCTCGACAAAGCCTCTTCGAGATTAACGGGATTAACACGGGAACT	1165	
Db	594	GGTCTATTGGGCAAGGAAGCCACTGGCTCGCGGATGACGGTTACAGCTAAGCATTCT	653	
Qy	1166	TCAGAAACCTCTTGGCCGCACGAGATGGAGCTTATTTATGTCAATCTTGAAGGCCCGCG	1225	
Db	654	GGAGAACCCGCTGGCCGCGGTGCAGATGGGGCTGTACTAGTGAACCCGAGGCGCGGAA	713	
Qy	1226	TGAAACACAGATCTCTTGCTTCCGCGAAGATATCAGGGAAGCTTTTTCAGCTATGCC	1285	
Db	714	CGGCAACCGGACCCCATGTCGCCGCGGGTCTGACATTCGGGAGACGTTTCGGCGCATGGC	773	
Qy	1286	CATGGATGATGAGGAGACTGTGCCCTTGATTCGCGGGAGGGCATACATTTGGTAAAGCACA	1345	
Db	774	CATGAACGAGCTCGAAACACGCGGCTGATCTGTCGGCGGTACACTTTCGTTAAGACCCA	833	
Qy	1346	TGTTGCAGGCTCTCTTGAAAAATGTATTGGCGCAGGGCCTGTATGGTGCACCTGTGAGGA	1405	
Db	834	TGGCGCGG----GCCCGCGCATCTGTTCGGCCCGCAACCGAGGCTGCTCCGCTGGAGCA	890	
Qy	1406	GCAGGACTGGGATGAAAAATAAATGTGTGTACAGGAACGGCAAAATATACCATACCAG	1465	
Db	891	GATGGCTTGGGCTGGNAGAGCTCGTATGSCACCGNACCGGTAAAGACCGGATACCAG	950	
Qy	1466	TGGCTTGGAAAGAGCCTGTGFCGA	1488	
Db	951	CGGCATCGAGGTCGTATGGACGA	973	

RESULT: 14
US-08-459-499-8

QY 928 CGGATACGTTATCTGGATAAGCCGCTGCGATCTGTGCGCCATCAAGAAAAATACG 987
Db 2385 CCGACAAACCCAGCTTGGCAAGGCGCGCGCTGTGCGCGTCAAGAAGAGTACG 2444
QY 988 GCTCCAGTATTCCTGGGAGACCTGATGGTCTCTGCTGCTAA--TGTTCCTTGAATC 1045
Db 2445 GCAAGAGCTCTATGGGCGGACCTGATGTTTTCGCGCGCAACCGCTGCGCTCGGAATC 2504
QY 1046 CATGGGATTAACACGCTGGGATTTGCTGGCGAAGAGAAGTACTGCTGGAGTGGACCT 1105
Db 2505 GATGGGCTTCAAGAGCTTGGGCTGGCTTCG--GCTGACAGTGGGAGACGATGA 2562
QY 1106 GGTATCTGGGCGCTGACACAGCCCTTTCGAGATAACCGGATAAACCGGAAACT 1165
Db 2563 GGTCTATTGGGCAAGGACCACTGCTGCGGATGACGCTTACAGCGTAAGCGATCT 2622
QY 1166 TCAGAAACCTCTGCGCCAGCAGATGGGACTTATTTATCTCAATCTGNAAGCCCGG 1225
Db 2623 CGAGAACCGCTGGCGCGGTGCGATGGGCTGATCTACGTGAACCGGAGCGCGAA 2682
QY 1226 TGGAAACAGATCTCTGGCTTCGCGAAGATATCAGGGAAGCTTTTTCACGTATGGC 1285
Db 2683 CGGCAACCGGACCCATGGCGCGGTGCGATTCGCGAGACGTTTCGGCGCATGGC 2742
QY 1286 CATGGATGATGAGAGACTGGCCCTGATCGCGGAGGCGATACATTTGGTAAAGCACA 1345
Db 2743 CATGAACAGCTGGAACAGCGGCGCTGATGCGCGGCTACACTTTCGTTAGACCCA 2802
QY 1346 TGGTCAGCGCTCTCTGAAAATGATTTGGCGCAGGCGCTGATGGTGCACCTGTGGAGGA 1405
Db 2803 TGGCGCG--GCCGCGCGATCTGGTCGCGCGCGAACCAGGCTCTCGCTGGAGCA 2859
QY 1406 CGAGGACTGGATGGAATAATGTTGTACAGGAACCGCAATATACCATCACCAG 1465
Db 2860 GATGGGCTTGGCTGGAAGAGCTGATGGCAGCGAAGCGGTAAAGACGCGATCACCAG 2919
QY 1466 TGGCTGGAAGGAGCCTGGTGA 1488
Db 2920 CGGCATCGAGTCTGATGGAGCA 2942

RESULT 15

US-08-313-185-45
; Sequence 45, Application US/08313185
; Patent No. 5851763

GENERAL INFORMATION:

; APPLICANT: Heym, Beate
; APPLICANT: Cole, Stewart
; APPLICANT: Young, Douglas
; APPLICANT: Zhang, Ying
; APPLICANT: Honore, Nadine
; APPLICANT: Telenti, Anello
; APPLICANT: Bodner, Thomas
; TITLE OF INVENTION: Rapid Detection of Antibiotic Resistance
; TITLE OF INVENTION: In Mycobacterium Tuberculosis
; NUMBER OF SEQUENCES: 66
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/313,185
; FILING DATE: 12-OCT-1994
; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Kenneth J.
; REGISTRATION NUMBER: 25,146
; REFERENCE/DOCKET NUMBER: 02356.0068-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 408-4000
; TELEFAX: (202) 408-4400
; INFORMATION FOR SEQ ID NO: 45:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4795 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-313-185-45

Query Match 18.4%; Score 274.2; DB 2; Length 4795;

Best Local Similarity 61.4%; Pred. No. 2.5e-79;

Matches 493; Conservative 0; Mismatches 303; Indels 7; Gaps 3;

QY	688	ATCCCTGGGGGCTGATTTGATATGCCACCAAGATTTCAACAGCTGGGATATGGAGCTC 747
Db	2145	ACCCGATGGGTGGCGGCTTCGACTATGCCGCGAGGTGCGACCTGACTTGCACGCC 2204
QY	748	TGAAAAAGATATCAAGATTTGCTGACAACTTCCAGGATTTGCTGCTCGGATATG 807
Db	2205	TGACGCGGACATCGAGGAAGTATGACCACTTCGACGCGGTGGTGGCCCGGACTAG 2264
QY	808	GTCATATGCTCTCTTTTATTCGATGGCTGGCACGGTCCGCGAATACAGSACAT 867
Db	2265	GCCACTACGGCGGCTGTTTATCCGATGGCGTGGCACCTCGCGCACCTACCCGATCC 2324
QY	868	ATGATGCGCGGAGCGCAGTGGTGGTACCAAGCTTTTGAACCGCTGAACAGCTGGC 927
Db	2325	ACGACGCGCGCGCGCGCGGCGGCGGATGATGCGGCTGCGCGCTTAACAGCTGGC 2384
QY	928	CGGATAACGTTAATCTGGATAAAGCCCGTGCATGCTGTGGCCAGTCAAGAAAAATACG 987
Db	2385	CGGACAAACCCAGCTTGGACAAAGCGCGCGCTGCTGTGGCGGCTCAAGAGAGTACG 2444
QY	988	GCTCCAGTATTCCTGGGAGACCTGATGGTCTCTGCTGCTGCTGCTGCTGCTGCTGCT 1045
Db	2445	GCAAGAGCTCTATGGGCGGACCTGATTTTTCGCGCGCAACCGCTGCGTTCGGAATC 2504
QY	1046	CATGGGATTAACACGCTGGGATTTGCTGGCGAAGAGAAGATGCTGGGATTCGACCT 1105
Db	2505	GATGGCTTCAAGACGTTGGGCTTCGCTCG--GCTGACCACTGGGAGACGATGA 2562
QY	1106	GGTATCTGGGCGCTGACACAGCCCTTTCGAGATAACCGGATAAACCGGAAACT 1165
Db	2563	GGTCTATTGGGCAAGGACCACTGCTGCGGATGACGCTTACAGCGTAAGCGATCT 2622
QY	1166	TCAGAAACCTCTGCGCCAGCAGATGGGACTTATTTATCTCAATCTGNAAGCCCGG 1225
Db	2623	CGAGAACCGCTGGCGCGGTGCGATGGGCTGATCTACGTGAACCGGAGCGCGAA 2682
QY	1226	TGGAAACAGATCTCTGGCTTCGCGAAGATATCAGGGAAGCTTTTTCACGTATGGC 1285
Db	2683	CGGCAACCGGACCCATGGCGCGGTGCGATTCGCGAGACGTTTCGGCGCATGGC 2742
QY	1286	CATGGATGATGAGAGACTGGCCCTGATCGCGGAGGCGATACATTTGGTAAAGCACA 1345
Db	2743	CATGAACAGCTGGAACAGCGGCGCTGATGCGCGGCTACACTTTCGTTAGACCCA 2802
QY	1346	TGGTCAGCGCTCTCTGAAAATGATTTGGCGCAGGCGCTGATGGTGCACCTGTGGAGGA 1405
Db	2803	TGGCGCG--GCCGCGCGATCTGGTCGCGCGCGAACCAGGCTCTCGCTGGAGCA 2859
QY	1406	CGAGGACTGGATGGAATAATGTTGTACAGGAACCGCAATATACCATCACCAG 1465
Db	2860	GATGGGCTTGGCTGGAAGAGCTGATGGCAGCGAAGCGGTAAAGACGCGATCACCAG 2919
QY	1466	TGGCTGGAAGGAGCCTGGTGA 1488

```

Db      2920 CGCATCGAGGTGCTATGGACGA 2942
|||||
RESULT 16
US-09-082-614A-45
; Sequence 45, Application US/09082614A
; Patent No. 6124098
; GENERAL INFORMATION:
; APPLICANT: Heym, Beate
; APPLICANT: Cole, Stewart
; APPLICANT: Young, Douglas
; APPLICANT: Zhang, Ying
; APPLICANT: Honore, Nadine
; APPLICANT: Telenti, Amalio
; APPLICANT: Bodmer, Thomas
; TITLE OF INVENTION: Rapid Detection of Antibiotic Resistance
; TITLE OF INVENTION: in Mycobacterium Tuberculosis
; NUMBER OF SEQUENCES: 66
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/082,614A
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/313,185
; FILING DATE: 12-OCT-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Kenneth J.
; REGISTRATION NUMBER: 25,146
; REFERENCE/DOCKET NUMBER: 02356.0068-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 408-4000
; TELEFAX: (202) 408-4400
; INFORMATION FOR SEQ ID NO: 45:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4795 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-09-082-614A-45
Query Match 18.4%; Score 274.2; DB 3; Length 4795;
Best Local Similarity 61.4%; Pred. No. 2.5e-79;
Matches 493; Conservative 0; Mismatches 303; Indels 7; Gaps 3;
QY      688 ATCCCTGGGGGCTGATTTTGGATATGCGCACCAGATTTCACAGCTGGATATGGAGCTC 747
Db      2145 ACCCGATGGGTGGGGGCTTCGACTATGCGCGGAGGTCGCGACAGTCGACTTGGCGCC 2204
QY      748 TGAAGAAATATCAAGATTGCTGACAACTTCCACAGGATTTGGTCCCTGCGGATATG 807
Db      2205 TGACGGGGACATCGAGGAAGTATGACCACTTCGACCGCTGGTGGCCGCGGACTACG 2264
QY      808 GTCAATATGTCCTTTCTTTATTCGTATGCTTGGCAGCGTGGCGGAACATACAGGACAT 867
Db      2265 GCCACTACGGCGCTGTTTATCCGGATGGCGTGGCAGCTGCGGCACTACCGCATCC 2324
QY      868 ATGATGCCGGGGAGGCGGCGGATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 927
|||||
2325 ACACGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2384
QY      928 CGGATAACGTTAATCTGGATAAAGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 987
Db      2385 CCACAAAGCCAGCTTGGACAAGCGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2444
QY      988 GCTCCAGTATTTCTGCGGGAGAGCTGATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1045
Db      2445 GCAAGAAGCTCTCATGCGCGGAGCTGATTTCTGCGCGGCAACCGCTGCGCTCGGAATC 2504
QY      1046 CATGGGATTTAAACACGCTGGATTGCTGCGCGAAGAGAGATGACTGGGAGTCGGACCT 1105
Db      2505 GATGGGCTTCAAGACGTTGCGGTTGCGCTTCGG--GCTCGACAGTGGGAGACCGGATGA 2562
QY      1106 GTTATATGGGGCGCTGACAAACAGCCTTTGCGAGATAACCGGGATAAAGCGGAAACT 1165
Db      2563 GGTCTATTGGGCAAGGAAGCCACCTGGCTCGCGGATGACGGTTACAGCGTAAGCGATCT 2622
QY      1166 TCAGAAACCTCTTTCGCGCCAGCAGATGGGACTTATTATGTCATCTCAATCTCAAGCCCGG 1225
Db      2623 GGAGAACCGCTGCGCGCGGTGCGAGATGGGGCTGATCTAGCTGAACCGGAGCGCGGAA 2682
QY      1226 TGAAGAACACAGATCCTCTGCTGCTCCGCGAAGAGATATCAGGAAAGCTTTTTCAGTATGGC 1285
Db      2683 CGCAACCGGACCCCATGCGCGCGGTGCGACATTCGCGAGACGTTTCGCGCATGGC 2742
QY      1286 CATGATGATGAGGAGACTGTGCGCTGATCGCGGAGGAGCATATTTGTTAAAGCACA 1345
Db      2743 CATGAACGAGCTCGAAACAGCGCGCTGATCGTGGCGGTGACACTTTTCGTTAAGACCCA 2802
QY      1346 TGGTGACGCTCTCTGAAATAATGTTATGCGCGAGGCGCTGTTGTTGTCACCTGTTGGAGGA 1405
Db      2803 TGGCGCG--GCCCGCGCGATCTGCTGCGCGCGGAAACCGGAGGCTGCTCGCTGGAGCA 2859
QY      1406 GCAGGACCTGGGATGGAATAATAATGCTGCTACAGGAAACCGGAAATATACCATCACAG 1465
Db      2860 GATGGGCTTGGCTGGAGAGCTGCTGATGACCGGAAACCGGAAACCGGATACCGATCACCAG 2919
QY      1466 TGGCCTGGAAGAGCGCTGTCGA 1488
Db      2920 CGCATCGAGGTGCTATGGACGA 2942

```

:	REGISTRATION NUMBER:	38,347			
:	REFERENCE/DOCKET NUMBER:	09015/002001			
:	TELECOMMUNICATION INFORMATION:				
:	TELEPHONE:	619/678-5070			
:	TELEFAX:	619/678-5099			
:	INFORMATION FOR SEQ ID NO: 5:				
:	SEQUENCE CHARACTERISTICS:				
:	LENGTH:	2262 base pairs			
:	TYPE:	nucleic acid			
:	STRANDEDNESS:	double			
:	TOPOLOGY:	linear			
:	MOLECULE TYPE:	cdna			
:	FEATURE:				
:	NAME/KEY:	Coding Sequence			
:	LOCATION:	1...2259			
:	US-08-674-887A-5				
<hr/>					
	Query Match	15.1%	Score	224.6;	DB 2; Length 2262;
	Best Local Similarity	57.4%;	Pred.	3.1e-63;	
	Matches	476;	Conservative	0;	Mismatches 329; Indels 24; Gaps 3;
<hr/>					
QY	682	AATCAAAATCCCTGGGGCGCTGATTTCATTATGCACACAGATTTCAACAGCTGTGATATGG	741		
DB	218	AATCAGACCCGATGGATCCGGATTTCACATACCCTGAAGAAGTACGCAAGCTCGATTTCG	277		
QY	742	AGGCTCTGAAAAAAGATATCAAAGATTTCGTGACAACACTTCCCAGAGTTGGTGCCCTCGG	801		
DB	278	ACGCCCTCGAAGAAAAGATGTCACACGCTTGATGACCGATAGCCAAGAGTGGTGGCCGCTG	337		
QY	802	ATTATGTGCTAATATGGCTCTTCCTTTTATTCGTATGGCTTCGACGGTGC CGGACATACA	861		
DB	338	ACTGGGGCACACTACGGCGGTTTGATGATCGTATGGCTTGGCATCTCGCTGGCACCTACC	397		
QY	862	GGACATATGATGGCCGGGAGGCCAGTGGTGGTCAGCAACACTTTTGAACCCGCTGAACA	921		
DB	398	GTATTGTGATGGCCGTGGGGCGGTGGTACC GGAAGCCAGCGCTTTGCACCGCTCACT	457		
QY	922	GCTGGCCGGATACGTTTAATCTGGATAAACCCGCTCGATTGCTGTGGCCAGTCAAGAAA	981		
DB	458	CTTGGCCGGACACGTCAGCGCTGGATAAAGCGGCGCGCTGCTGTGGCCGATCAAGAAGA	517		
QY	982	AATACGGCTCCAGTATTTCCTGGGGAGACCTGATGTGCTGACTGGTAAATGTTGCCCTTG	1041		
DB	518	AGTACGGCAACAAAATCAGCTGGGACACCTGATGATCTGGCTGGCACCGTGGCTTATG	577		
QY	1042	AATCCATGGATTAAACACCTGGGATTTCTCTGGCGGAAGAAGATGACTGGGAGTCGG	1101		
DB	578	AGTCCATGGGCTTACCTGCTTAGCGGCTCTCTTTTCGGCCGCGTCGATATTGGGAACCG	637		
QY	1102	ACCTGG---TATACTGGGGCCGTGACAAACAGCCTCTTCGAGAT-----	1142		
DB	638	AAAAGATATCTACTGGGGTGACAAAAGAGTGGCTGGCACCTTCTGACGAACGCTACG	697		
QY	1143	--AACCGGGATAAAAAACGGGAAACTTTCAGAAACCTTTGGCCGACGACGATGGGACTTA	1200		
DB	698	CGCAGCTGAAACAGCCAGAGACCATGMAAACCCGCTGGCGCTGTCCAATGGGTCTGA	757		
QY	1201	TTTTATCTCAATCCTGAAGCCCCGGTGGAAAACCATCTCTGGGCTTCGGGAAAGATA	1260		
DB	758	TCATATGTGAACCCGGAAGTGTTTAACGGCCACCTGATCGCTGAGAAGCCGACAGCAG	817		
QY	1261	TCAGGGAAGCTTTTTTCACGCTATGGCCATGGATGATGAGGAGACTGTGGCCCTGATCGCG	1320		
DB	818	TACTTGAACCTTCGCCCGTATGGCGATGAACGACGAAAAAACCGCAGCCCTCACAGCTG	877		
QY	1321	GAGGCGATACATTGTTGAAGCACATPGTGGCAGCGCTCTCTTGA AAAATATATTGGCGCAG	1380		
DB	878	CGCGCCACACCGCTCGGTAATTGTACGGGTAATGGCAATGCCCTCTGCGT---TAGCCCTG	934		
QY	1381	GGCCTCATGTGCACCTCTGGAGGACACGGGACTGGGATGGAAAAATAAATGTGTGTACAG	1440		
DB	935	ACCCAAAGCCCTTGACCTTGAAACACCGGGCTTAGTTGGGGCAACCCCAACATGCGAG	994		

US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZgpt-F1s
US-08-232-463-14

Query Match 3.2%; Score 48; DB 1; Length 7218;
Best Local Similarity 3.1%; Pred. No. 5.6e-05;
Matches 12; Conservative 217; Mismatches 157; Indels 0; Gaps 0;
QY 1065 GGATTTCGTCGGGAGAGAGATGACTGGGAGTCGGACCTGGTATCTACTGGGGCCCTGAC 1124
DB 1445 GAAATTTGCTACRR 1386
QY 1125 AACAGCCTCTGCAGATAACCGGGATAAACCGGAACTTCAGAACTCTTGCGGCC 1184
DB 1385 RRR 1326
QY 1185 ACGCAGATGGGACTATTATGTCAATCTCAAGCCCGCTGAGAAACAGATCCCTCG 1244
DB 1325 RRR 1266
QY 1245 GCTCCCGGAAGATATCAGGAAGCTTTTTCAGTATGGCCATGGATGATGAGAGACT 1304
DB 1265 RRR 1206
QY 1305 GTGGCCCTGATCGGGGAGGCATACATTGGTAAACACATGCTGAGCGCTCTCTGAA 1364
DB 1205 RRR 1146
QY 1365 AAATGTATTGGCGCAGGCGCTGATGTGCACCTGTGAGGAGCAGGACTGGGATGCAA 1424

Db 1145 RRR 1086
QY 1425 AATAAATGCTGTACAGGAACGGCAA 1450
DB 1085 RRR 1060

RESULT 21
US-09-125-287-2
; Sequence 2, Application US/09125287B
; Patent No. 6114602
; GENERAL INFORMATION:
; APPLICANT: BARG, Rivka
; APPLICANT: SALTS, Tehiam
; TITLE OF INVENTION: METHOD FOR THE INTRODUCTION OF GENETIC PARTHENOCAOPY IN
; TITLE OF INVENTION: PLANTS
; FILE REFERENCE: INTRO GENETIC PARTHENOCAOPY IN PLANTS
; CURRENT APPLICATION NUMBER: US/09/125,287B
; CURRENT FILING DATE: 1998-11-09
; EARLIER APPLICATION NUMBER: PCT/IL97/00051
; EARLIER FILING DATE: 1997-02-13
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 2
; LENGTH: 4518
; TYPE: DNA
; ORGANISM: TPRP-F1 PROMOTOR
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (6)..(191)
; OTHER INFORMATION: "n"s are any nucleic residue
US-09-125-287-2

Query Match 2.68; Score 38.8; DB 3; Length 4518;
Best Local Similarity 43.8%; Pred. No. 0.045;
Matches 169; Conservative 0; Mismatches 217; Indels 0; Gaps 0;
QY 424 CTTAGTATTTTGTCTGTAGTATCTATCCAGCAATAGGTATATCTGTGTCATCAATA 483
DB 3826 CTTATTGATGATTGTGGCAAGTATCCACCACCAAGACTTCAAGTGTCTTCAACTAATC 3885
QY 484 AAGTTGACTTTTGTATACACATCGGAATTCCTTAATCCGGAGCTATTTCGATGATAA 543
DB 3886 AAGGTTTCCACTTTTCTCACTTTCATATTACTCCTACTCAATTTATGTGTGTAC 3945
QY 544 AAAAAACTCTCTCTGATCTTCTGCGCTATCGGGAGCTTTTCTACCGCTGTAG 603
DB 3946 AGTTGACATTTCAAGTATTAGGCCCAATTTCTTAGCTCGGAATTTTAAATCTCTTT 4005
QY 604 CCGCTGATAAAAAGAGACTCAAAATTTCTACTATCCAGAAACACTGGATTAACTCCTC 663
DB 4006 AAATATTTTGTATTATCTACTTATTACGTAGTTTTCATAAGTATAAATTCATTTCATA 4065
QY 664 TGAGATTACACACCCCTGAATCAATCCCTCGGGGCTGATTTGATTTGACACAGAT 723
DB 4066 TATGAATTCACGCTCAAAAATTTAAAGTTTATTAAACCAATGCCACATAACTTAGAACAT 4125
QY 724 TTCAACAGCTGGGATATGGAGGCTCTCAAAAAGATATCAAGATTTTGTGACACTCCC 783
DB 4126 ACAATATCTTTTGTATCAAGATTTGGCAATTCGTATACAATAATCTTTAGCAAGTAATA 4185
QY 784 AGGATTGGTCCCTCGGGGATTTATGTT 809
DB 4186 TGTATACCAACATTTATGTAATGAT 4211

RESULT 22
US-09-125-287-1
; Sequence 1, Application US/09125287B
; Patent No. 6114602
; GENERAL INFORMATION:
; APPLICANT: BARG, Rivka

```
; APPLICANT: SALTS, Tehiam
; TITLE OF INVENTION: METHOD FOR THE INTRODUCTION OF GENETIC PARTHENOCAOPY IN
; FILE REFERENCE: INTRIO GENETIC PARTHENOCAPI IN PLANTS
; CURRENT APPLICATION NUMBER: US/09/125,287B
; CURRENT FILING DATE: 1998-11-09
; EARLIER APPLICATION NUMBER: PCT/IL97/00051
; EARLIER FILING DATE: 1997-02-13
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 12839
; TYPE: DNA
; ORGANISM: TPRP-F1 GENOMIC CLONE
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (5)..(11322)
; OTHER INFORMATION: "n"'s are any nucleic residue
US-09-125-287-1

Query Match      2.6%; Score 38.8; DB 3; Length 12839;
Best Local Similarity 43.8%; Pred. No. 0.083;
Matches 169; Conservative 0; Mismatches 217; Indels 0; Gaps 0;

QY 424 CTTAGTATTTTGTCTGTAGTATCTATCCAGCAATAGGTATATCTCTGTGCAATCAATA 483
Db 3827 CTTATTGATGATTTGTGGCAAGTATCCACCAAGACTTCAAGTGTCTTCAACCTAAATC 3886
QY 484 AAGTGCATTTTCTATACAACTGCGAATTTTCCCTTAATCCGGAGCTATTTCGATGATAA 543
Db 3887 AAGTTTCCACTTTTCTCACTTTCAATTAATTAATCACTCACTCACTCACTCACTCACT 3946
QY 544 AAAAACTCTCTGTTCTGTCTGTTCTGCGCTATCGGGAGCTTTTCTACCGCTGTAG 603
Db 3947 AGTTGACATTTCAAGTATTAGGCCCAATTTTCTAGCTCGAATTTTTTTTAAATCTCTTT 4006
QY 604 CCCTGATATAAAGAGACTCAAAATTTCTACTATCCAGAAACACTGGATTAACTCCCTC 663
Db 4007 AATATTTTGATTTATATCTACTTTTATACGTAGTTTTCATAAGTATAAATTTTCATTTCA 4066
QY 664 TGAGATTACACAGCCCTGAATCAATCAATCCCTGGGGGCTGATTTTGTATGATCCACCAT 723
Db 4067 TATGAATTCAGCTCAAAATTTTAAAGTTTATTATCAATGCCATCAATAGAACAT 4126
QY 724 TTCAACAGCTGGATATGGAGGCTCTGAAAAAGATATCAAGATTTCGTGACAACTTCCC 783
Db 4127 ACAATATCTTTTGTATCAAGATTGCGAATTCGTATACATAATCTTTAGCAAGTAATA 4186
QY 784 AGATTGGTGGCTCGCGGATATGTT 809
Db 4187 TGTATACCAACATTATGTAATATGAT 4212

RESULT 23
US-08-991-789A-147/c
; Sequence 147, Application US/08991789A
; Patent No. 6225054
; GENERAL INFORMATION:
; APPLICANT: Frudakis, Tony N.
; REED, Steven G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; TREATMENT AND DIAGNOSIS OF BREAST CANCER
; NUMBER OF SEQUENCES: 292
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED IP Law Group
; STREET: 701 Fifth Avenue, Suite 6300
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/062,451
; FILING DATE: 04-APR-1997
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.419C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 147:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 579 base pairs
; MEDIUM TYPE: Floppy disk
```

```
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/991,789A
; FILING DATE: 11-Dec-1997
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Potter, Jane E. R.
; REGISTRATION NUMBER: 33,332
; REFERENCE/DOCKET NUMBER: 210121.419C3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 147:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 579 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 147:
US-08-991-789A-147

Query Match      2.3%; Score 33.8; DB 4; Length 579;
Best Local Similarity 57.8%; Pred. No. 0.57;
Matches 59; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

QY 1316 CGCGGGAGGCGATACATTTGGTAAAGCACATGTCGAGCGTCTCCTGAAAAATGTATGG 1375
Db 427 CGGTGGAGCCATGATGAGGCGACACACGCGGGGAGCTGGTCTGTCAGAACGGCTTTG 368
QY 1376 CGCAGGGCCCTGATGGTGCACCTGTGGAGGAGCAGGAGCTGGG 1417
Db 367 CTGACATCATCCCGTGTACCAGCAGAAAGTGGAGGACGTTGG 326

RESULT 24
US-09-062-451-147/c
; Sequence 147, Application US/09062451
; Patent No. 6344550
; GENERAL INFORMATION:
; APPLICANT: Frudakis, Tony N.
; APPLICANT: Reed, Steven G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; TREATMENT AND DIAGNOSIS OF BREAST CANCER
; NUMBER OF SEQUENCES: 297
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED AND BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/062,451
; FILING DATE: 04-APR-1997
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.419C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 147:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 579 base pairs
```

```
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-09-062-451-147

Query Match      2.3%; Score 33.8; DB 4; Length 579;
Best Local Similarity 57.8%; Pred. No. 0.57;
Matches 59; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

Qy 1316 CGCGGGAGGCGATACATTTGGTAAAGCAGCATGTCAGCGTCTCCTGAAAAATGTATTGG 1375
   || |||| || || |||| || || |||| || || |||| || || |||| || || ||||
Db 427 CGGTGGAGCCAGTGAGATGGACACACACACGGGCGAGCTGTCNTGCGAAGCGCTTGG 368

Qy 1376 CGCAGGCGCTGATGGTGACCTGTGTGGAGGAGCAGGACTGGG 1417
   | | | | | | | | | | | | | | | | | | | | | | | |
Db 367 CTGACATCATCACCGTGTACACGACAGAGGTGGAGGACGTGG 326

RESULT 25
US-09-598-326-147/C
; Sequence 147, Application US/09598326
; Patent No. 6423496
; GENERAL INFORMATION:
; APPLICANT: Frudakis, Tony N.
;              Smith, John M.
;              Reed, Steven G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
;                     TREATMENT AND DIAGNOSIS OF BREAST CANCER
; NUMBER OF SEQUENCES: 247
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seed Intellectual Property Law Group PLLC
; STREET: 701 Fifth Avenue, Suite 6300
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/598,326
; FILING DATE: 20-Jun-2000
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Potter, Jane E.R.
; REGISTRATION NUMBER: 33,332
; REFERENCE/DOCKET NUMBER: 210121.419D1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 147:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 579 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 147:

US-09-598-326-147

Query Match      2.3%; Score 33.8; DB 4; Length 579;
Best Local Similarity 57.8%; Pred. No. 0.57;
Matches 59; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

Qy 1316 CGCGGGAGGCGATACATTTGGTAAAGCAGCATGTCAGCGTCTCCTGAAAAATGTATTGG 1375
   || |||| || || |||| || || |||| || || |||| || || |||| || || ||||
Db 427 CGGTGGAGCCAGTGAGATGGACACACACACGGGCGAGCTGTCNTGCGAAGCGCTTGG 368

Qy 1376 CGCAGGCGCTGATGGTGACCTGTGTGGAGGAGCAGGACTGGG 1417
   | | | | | | | | | | | | | | | | | | | | | | | |
Db 367 CTGACATCATCACCGTGTACACGACAGAGGTGGAGGACGTGG 326
```

```
RESULT 26
US-08-716-351A-5/C
; Sequence 5, Application US/08716351A
; Patent No. 6033905
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Gibbon Ape Leukemia Virus-Based
; TITLE OF INVENTION: Retroviral Vectors
; NUMBER OF SEQUENCES: 5
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/716,351A
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/03784
; FILING DATE: 06-APR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian,, Kevin L.
; REGISTRATION NUMBER: 34,774
; REFERENCE/DOCKET NUMBER: 15280-128-1PC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 543-9600
; TELEFAX: (415) 543-5043
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10970 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 1..10970
; OTHER INFORMATION: /standard_name= "p537 retroviral
; vector"
US-08-716-351A-5

Query Match      2.3%; Score 33.8; DB 3; Length 10970;
Best Local Similarity 51.7%; Pred. No. 3.3;
Matches 77; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

Qy 1267 AAGCTTTTTCACGTATGGCCATGGATGATGAGGAGACTGTGCCCTGTGATCGCGGGAGGC 1326
   || |||| || || |||| || || |||| || || |||| || || |||| || || ||||
Db 6662 AAGGTTTCTCAGATGGTAAATAGATGCGGGGGATCAGCTGGACCTGGACACAGAAGTCT 6603

Qy 1327 ATACATTTGTTAAAGCACATGGTGACGCTCTCCTGAAAAATGTATTGGCGGAGGCCCTG 1386
   || |||| || || |||| || || |||| || || |||| || || |||| || || ||||
Db 6602 TTAGACTGATTAATAACTAGGTGGAGAGGAGGGGTGAGGCCACTGTGTCAGGCCCCAC 6543

Qy 1387 ATGGTGACCTGTGGAGGAGCAGGACTG 1415
   || |||| || || |||| || || |||| || || |||| || || |||| || || ||||
Db 6542 CAGCTATGTTTTCAGGGGAGCAGACTG 6514

RESULT 27
US-09-134-001C-2054
; Sequence 2054, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCC
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
```


FEATURE: CDS
NAME/KEY: 14046..20036
LOCATION: 14046..20036
FEATURE: CDS
NAME/KEY: 20110..31284
LOCATION: 20110..31284
FEATURE: CDS
NAME/KEY: 31329..36071
LOCATION: 31329..36071
FEATURE: CDS
NAME/KEY: 36155..41830
LOCATION: 36155..41830
US-08-804-227C-7

Query Match 2.2%; Score 32.2; DB 2; Length 44377;
Best Local Similarity 53.6%; Pred. No. 26;
Matches 67; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY 167 GGAAGCATCAGAGGGCGGGAAGTCCGTCGGCGCAGTGAACCGTGCACACTCCGGG 226
DB 31666 GAGATCCCAAGAGGGCGGCTGGAACCGTGCAGCGCCCTGCACGAAGCGCCCTCGCGG 31607
QY 227 CAGTACATCCCGCGCGCTGATACCGGCAAGAATGGTCGCAAACTCCCGCTCCGTGCAG 286
DB 31606 CAGTACGAGTCCCGGCGCTGTCGGGGTCCGGCGAGTACCTTCCCTCCAGTCCCGCGG 31547
QY 287 CGGGC 291
DB 31546 CGGTC 31542

RESULT 33
US-08-804-198-1/c
; Sequence 1, Application US/08804198
; Patent No. 5945320
; GENERAL INFORMATION:
; APPLICANT: Burgett, Stanley G.
; APPLICANT: Kubstoss, Stuart A.
; APPLICANT: Rao, Nagaraja R.
; APPLICANT: Richardson, Mark A.
; APPLICANT: Rostock, Paul R., Jr.
; TITLE OF INVENTION: PLATENOLIDE SYNTHASE GENE
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PAUL R. CANTRELL 1138
; STREET: LILLY CORPORATE CENTER
; CITY: INDIANAPOLIS
; STATE: IN
; COUNTRY: USA
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: Macintosh 7.0
; SOFTWARE: Microsoft Word 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/804,198
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: CANTRELL, PAUL R.
; REGISTRATION NUMBER: 36,470
; REFERENCE/DOCKET NUMBER: P9113
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 317-276-3885
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 44377 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:

NAME/KEY: CDS
LOCATION: 350..14002
FEATURE: CDS
NAME/KEY: 14046..20036
LOCATION: 14046..20036
FEATURE: CDS
NAME/KEY: 20110..31284
LOCATION: 20110..31284
FEATURE: CDS
NAME/KEY: 31329..36071
LOCATION: 31329..36071
FEATURE: CDS
NAME/KEY: 36155..41830
LOCATION: 36155..41830
US-08-804-198-1

Query Match 2.2%; Score 32.2; DB 2; Length 44377;
Best Local Similarity 53.6%; Pred. No. 26;
Matches 67; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY 167 GGAAGCATCAGAGGGCGGGAAGTCCGTCGGCGCAGTGAACCGTGCACACTCCGGG 226
DB 31666 GAGATCCCAAGAGGGCGGCTGGAACCGTGCAGCGCCCTGCACGAAGCGCCCTCGCGG 31607
QY 227 CAGTACATCCCGCGCGCTGATACCGGCAAGAATGGTCGCAAACTCCCGCTCCGTGCAG 286
DB 31606 CAGTACGAGTCCCGGCGCTGTCGGGGTCCGGCGAGTACCTTCCCTCCAGTCCCGCGG 31547
QY 287 CGGGC 291
DB 31546 CGGTC 31542

RESULT 34
US-09-156-316-11/c
; Sequence 11, Application US/09156316
; Patent No. 6183961
; GENERAL INFORMATION:
; APPLICANT: Bernstein, Harold S.
; APPLICANT: Coughlin, Shaun R.
; TITLE OF INVENTION: Methods and Compositions for Regulating Cell Cycle
; TITLE OF INVENTION: Progression
; FILE REFERENCE: UCSF-020/01US
; CURRENT APPLICATION NUMBER: US/09/156,316
; CURRENT FILING DATE: 1998-09-18
; EARLIER APPLICATION NUMBER: 60/060,688
; EARLIER FILING DATE: 1997-09-22
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 2837
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-156-316-11

Query Match 2.1%; Score 32; DB 4; Length 2837;
Best Local Similarity 50.0%; Pred. No. 5.7;
Matches 106; Conservative 0; Mismatches 105; Indels 1; Gaps 1;

QY 423 CCTAGTATTTTGTGCTGTAGTATCTATCCAGCAATAGGTATATCTCTGTTGCATCA-A 481
DB 2581 CATCAGTTTCAGCCTTCTAGAGTATGAAACCGGCAATTAATTAATCTCTGTGACAGATA 2522
QY 482 TAAAGTTGACTTTTGTATACACATCGGAATTTCCCTTAATCCGGAGCTATTCTGTATGAT 541
DB 2521 TAACTGTACTTCAGATTTTGTACTTAAGTCTCTTTCTCCAGCAGCAATCAGCATAT 2462
QY 542 AAAAAAACTCTTCCTGTTCTGATCTCTTGGCGGTATCGGGAGGCTTTTCTACCGCTGT 601
DB 2461 CTATGTTGAAGTTCCCTTTCTCTTTCTTTGTTGCGTGAACGCTCTTTCTTTAGACACTCT 2402
QY 602 AGCCGCTGATAAAAAGAGACATCAAAATTTCT 633

Db 2401 AGCTCCGGGGAATAGCAGAATCTTCATGTTT 2370

RESULT 35

US-09-442-055-1

; Sequence 1, Application US/09442055
; Patent No. 6306631
; GENERAL INFORMATION:
; APPLICANT: Andersen, Jens Tonne
; APPLICANT: Ehrlich, Stanislas Dusko
; TITLE OF INVENTION: No. 6306631el ppGpp Synthetase and Expression
; FILE REFERENCE: 5259.204-US
; CURRENT APPLICATION NUMBER: US/09/442,055
; CURRENT FILING DATE: 1999-11-16
; EARLIER APPLICATION NUMBER: 0726/97
; EARLIER FILING DATE: 1997-06-20
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 2278
; TYPE: DNA
; ORGANISM: rela-bac
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (21)...(2222)
US-09-442-055-1

Query Match 2.1%; Score 31.8; DB 4; Length 2278;
Best Local Similarity 57.6%; Pred. No. 5.8;
Matches 57; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY 292 TATTTCAGGATACCTTCGTCATCAACACGCTACAAACAGAACACGCTTTTGTCT 351
DB 618 TATTACAGATGTCAACCTCATGAAGAAGAACGTCGACAGAGCTTTATGTCAT 677
QY 352 GACATCCCAAGAGAGGAAATATTCAGTCTGCCACCA 390
DB 678 GAGGTGTCAATGAAGTGAAGAAGAGTGTGCGAAGAAGTA 716

RESULT 36

US-09-282-147-38

; Sequence 38, Application US/09282147
; Patent No. 6274147
; GENERAL INFORMATION:
; APPLICANT: VAKHARIA, Vikram
; APPLICANT: YAO, Kun
; TITLE OF INVENTION: METHOD FOR GENERATING NONPATHOGENIC, INFECTIOUS
; TITLE OF INVENTION: PANCREATIC NECROSIS VIRUS (IPNV) FROM SYNTHETIC RNA
; FILE REFERENCE: 8288-9023
; CURRENT APPLICATION NUMBER: US/09/282,147
; CURRENT FILING DATE: 1999-03-31
; EARLIER APPLICATION NUMBER: US/60/080,278
; EARLIER FILING DATE: 1998-03-31
; EARLIER APPLICATION NUMBER: PCT/US97/12955
; EARLIER FILING DATE: 1998-03-31
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 38
; LENGTH: 3097
; TYPE: DNA
; ORGANISM: Infectious pancreatic necrosis virus
US-09-282-147-38

Query Match 2.1%; Score 31.8; DB 4; Length 3097;
Best Local Similarity 59.3%; Pred. No. 7;
Matches 54; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

QY 1050 GGATTTAAACGCTGGGATTTGTCGCGGAAGAGAAGATGACTGGAGTCGGACCTGGTA 1109
DB 2060 GGAGTTCCCAAGGATGTTATTCACCGCGAAATCGCGGATGACGAGGAGACATACCA 2119

QY 1110 TACTGGGGCGCTGACAAACAGCCTCTTTCGAG 1140
DB 2120 ATTCGGGTGTAGACATCAAGCCATCGCAG 2150

RESULT 37

US-09-196-390-5
; Sequence 5, Application US/09196390
; Patent No. 6307125
; GENERAL INFORMATION:
; APPLICANT: Block, Martina
; APPLICANT: Lorz, Horst
; APPLICANT: Lutticke, Stephanie
; APPLICANT: Walter, Lennart
; APPLICANT: Froberg, Claus
; APPLICANT: Kossmann, Jens
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING ENZYMES
; TITLE OF INVENTION: FROM WHEAT WHICH ARE INVOLVED IN STARCH
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: James F. Haley, Jr., c/o Fish & Neave
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10020
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/196,390
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 196 21 588.9
; FILING DATE: 29-MAY-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 196 36 917.7
; FILING DATE: 11-SEP-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP97/02793
; FILING DATE: 28-MAY-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Haley, Jr., James F.
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: AGREVO-9
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 596-9000
; TELEFAX: (212) 596-9090
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2825 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA to mRNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Triticum aestivum L.
; STRAIN: cv. Florida
; TISSUE TYPE: ca. 21 d Caryopses
; IMMEDIATE SOURCE:
; LIBRARY: CDNA library in pBluescript sk (-)
; CLONE: pTASS1
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 162...2559
US-09-196-390-5

Db 2852 TCTCCACACAGACATCACAGGCTGCTTGAAGACCTTCTTGCCCGAGGCTCTGA 2795

RESULT 40

US-08-365-486A-20/c
; Sequence 20, Application US/08365486A
; Patent No. 5834306
; GENERAL INFORMATION:
; APPLICANT: Webster, Keith A.
; APPLICANT: Bishopric, Nanette H.
; TITLE OF INVENTION: Tissue Specific Hypoxia Regulated
; TITLE OF INVENTION: Therapeutic Constructs
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Avenue, Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/365,486A
; FILING DATE: 23-DEC-1994
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Sholtz, Charles K.
; REGISTRATION NUMBER: 38,615
; REFERENCE/DOCKET NUMBER: 8255-0018
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-0880
; TELEFAX: (415) 324-0960
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4780 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: Human NOS-SN gene, Nakane, et al,
; INDIVIDUAL ISOLATE: FEBS Lett 316:175 (1993)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 431..4732
US-08-365-486A-20

Query Match 2.1%; Score 31.6; DB 2; Length 4780;
Best Local Similarity 54.2%; Pred. No. 11;
Matches 64; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

QY 1270 CTTTTCACGTATGCCATGATGATGAGGAGACTGTGGCCCTGATCGCGGAGGGCATA 1329
Db 3339 CTCCTCCAGCTGCGATCATCTGATGAGGGAATTGTGGCCCTTTTCAATGTTGACATCA 3280
QY 1330 CATTGGTAAAGCACATGGTGACGCTCTCTGAAAAATGTTATGGCGAGGCGCTGA 1387
Db 3279 TCTCCACACAGACATCACAGGCTGCTTGNAGACCTTCTTGCCCGAGGCTCTGA 3222

RESULT 41

US-09-123-708-3/c
; Sequence 3, Application US/09123708
; Patent No. 6146887
; GENERAL INFORMATION:
; APPLICANT: SCHRADER, Juergen
; APPLICANT: GODECKE, Axel

; TITLE OF INVENTION: DNA EXPRESSION VECTORS FOR USE IN GENE THERAPEUTIC
; TITLE OF INVENTION: TREATMENT OF VASCULAR DISORDERS
; FILE REFERENCE: 511169-2003
; CURRENT APPLICATION NUMBER: US/09/123,708
; CURRENT FILING DATE: 1998-07-28
; EARLIER APPLICATION NUMBER: 08/553,503
; EARLIER FILING DATE: 1996-03-01
; EARLIER APPLICATION NUMBER: P4411402.8
; EARLIER FILING DATE: 1994-03-31
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 3
; LENGTH: 4780
; TYPE: DNA
; ORGANISM: Cytomegalovirus
US-09-123-708-3

Query Match 2.1%; Score 31.6; DB 3; Length 4780;
Best Local Similarity 54.2%; Pred. No. 11;
Matches 64; Conservative 0; Mismatches 54; Indels 0; Gaps 0;
QY 1270 CTTTTCACGTATGCCATGATGATGAGGAGACTGTGGCCCTGATCGCGGAGGGCATA 1329
Db 3339 CTCCTCCAGCTGCGATCATCTGATGAGGGAATTGTGGCCCTTTTCAATGTTGACATCA 3280
QY 1330 CATTGGTAAAGCACATGGTGACGCTCTCTGAAAAATGTTATGGCGAGGCGCTGA 1387
Db 3279 TCTCCACACAGACATCACAGGCTGCTTGNAGACCTTCTTGCCCGAGGCTCTGA 3222

RESULT 42

US-09-123-624-3/c
; Sequence 3, Application US/09123624
; Patent No. 6149936
; GENERAL INFORMATION:
; APPLICANT: SCHRADER, Juergen
; APPLICANT: GODECKE, Axel
; TITLE OF INVENTION: DNA EXPRESSION VECTORS FOR USE IN THE GENE THERAPEUTIC
; TITLE OF INVENTION: TREATMENT OF VASCULAR DISORDERS
; FILE REFERENCE: 511169-2004
; CURRENT APPLICATION NUMBER: US/09/123,624
; CURRENT FILING DATE: 1998-07-28
; PRIOR APPLICATION NUMBER: 08/553,503
; PRIOR FILING DATE: 1996-03-01
; PRIOR APPLICATION NUMBER: 4411402.8
; PRIOR FILING DATE: 1994-03-31
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 3
; LENGTH: 4780
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-123-624-3

Query Match 2.1%; Score 31.6; DB 3; Length 4780;
Best Local Similarity 54.2%; Pred. No. 11;
Matches 64; Conservative 0; Mismatches 54; Indels 0; Gaps 0;
QY 1270 CTTTTCACGTATGCCATGATGATGAGGAGACTGTGGCCCTGATCGCGGAGGGCATA 1329
Db 3339 CTCCTCCAGCTGCGATCATCTGATGAGGGAATTGTGGCCCTTTTCAATGTTGACATCA 3280
QY 1330 CATTGGTAAAGCACATGGTGACGCTCTCTGAAAAATGTTATGGCGAGGCGCTGA 1387
Db 3279 TCTCCACACACAGACATCACAGGCTGCTTGAAGACCTTCTTGCCCGAGGCTCTGA 3222

RESULT 43

US-08-880-342-20/c
; Sequence 20, Application US/08880342
; Patent No. 6218179
; GENERAL INFORMATION:
; APPLICANT: Webster, Keith A.

RESULT 45

```

US-09-007-005-17
; Sequence 17, Application US/09007005B
; Patent No. 6258558
; GENERAL INFORMATION:
; APPLICANT: Szostak, Jack W.
; APPLICANT: Roberts, Richard W.
; APPLICANT: Liu, Rihc
; TITLE OF INVENTION: FUSIONS
; FILE REFERENCE: 00786/350003
; CURRENT APPLICATION NUMBER: US/09/007,005B
; CURRENT FILING DATE: 1998-01-14
; EARLIER APPLICATION NUMBER: 60/035,963
; EARLIER FILING DATE: 1997-01-27
; EARLIER APPLICATION NUMBER: 60/064,491
; EARLIER FILING DATE: 1997-11-06
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 289
; TYPE: RNA

```

RESULT 44
US-08-253-155A-7
; Sequence 7, Application US/08253155A
; Patent No. 5691147
; GENERAL INFORMATION:
; APPLICANT: Gyrulis, Jeno
; APPLICANT: Draetta, Giulio


```

; CURRENT FILING DATE: 1999-07-20
; EARLIER APPLICATION NUMBER: PCI/EP98/00388
; EARLIER FILING DATE: 1998-01-19
; NUMBER OF SEQ ID NOS: 9
; SEQ ID NO 5
; LENGTH: 4160
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: MAKAR17
; LOCATION: 945..968
; OTHER INFORMATION: region corresponding to
; FEATURE:
; NAME/KEY:
; LOCATION: 1444..1445
; OTHER INFORMATION: insertion point of T-D
; OTHER INFORMATION: insertion in the line
; FEATURE:
; NAME/KEY:
; LOCATION: 1418..1570
; OTHER INFORMATION: putative open reading
; US-09-341-678-5

```

	Query Match	2.1%	Score 30.8;	DB 4;	Length 4160;
	Best Local Similarity	55.7%;	Pred. No. 18;		
	Matches	59;	Conservative 0;	Mismatches 47;	Indels 0; Gaps 0;
Qy	736	ATATGGAGGCTCTGAAAAAGATATCAAGAGTTTGTGACAACATCCACAGATTGGGCC	795		
Db	4035	ATATCTTGGCTTTGGAAATCAAAAGTGAGAGTTATGCAGACAATGGTTCAGGTTTGACCG	3976		
Qy	796	CTGGGATTAATGGTCATATATGGTCCCTTTCTTTATTCGTATGGCTTG	841		
Db	3975	AAACTGACTCTGTTTAAAAACCTTCATATATTTTTTCTGATATCTTG	3930		

```

RESULT 49
US-09-103-840A-2
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

Query Match                2.1%; Score 30.8; DB 4; Length 4403765;
Best Local Similarity      57.1%; Pred.No. 67;
Matches 56; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

Qy      887  CAGTGGTGGTGCAGCAACGTTTGAACCGCTGACAGCTGCCGGATACCGTTAATCTGGA 946
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      4113218  CAGTGCAGGTGCGGACGATCGGGCGGTACACACTGTACCGGATCCGGCTATTGGGC 4113277

Qy      947  TAAAGCCCGTCGATTCGTGTGGCCAGTCAAGAAAAAT 984
          ||||| || ||||| || ||||| ||||| ||||| ||||| ||||| |||||
Db      4113278  ACAAGCCAGTTCCTATTGGGCACGAAGCAATTAGAAT 4113315

```

```

RESULT 50
US-09-103-840A-1
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen K.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv
US-09-103-840A-1

Query Match          2.1%; Score 30.8; DB 4; Length 4411529;
Best Local Similarity 57.1%; Pred. No. 67;
Matches 56; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY      887  CAGTGGTGGTTCAGCAACGTTTTTGAACCGCTGAACAGCTGCCCGGATAACGTTAACTCTGGA 946
         ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB      4121023 CAGTGCAGGTCTGGGACAGTCGGGGCGGTACACACTCTGCACGGGATCCGGCTATTGGGC 4121082
         ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY      947  TAAAGCCCGTCGATTGCTGTGGCCAGTCAAGAAAAAT 984
         ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB      4121083 ACAAGCCAGTTCTCTATTGGGCACAAGCCAAATTAGAAT 4121120
         ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

Search completed: June 19, 2003, 03:36:03
Job time : 91.1139 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 18, 2003, 01:20:59 ; Search time 1574.88 Seconds

(without alignments)
15312.305 Million cell updates/sec.

Title: US-09-674-277-1

Perfect score: 1489

Sequence: 1 ctgcagtcgagatgaaag.....ctggaaggagcctggtgcac 1489

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 150 summaries

Database :

EST:*

1: em_estba:*

2: em_esthum:*

3: em_estin:*

4: em_estmu:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_htc:*

9: gb_estl:*

10: gb_est2:*

11: gb_htc:*

12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em_estfun:*

16: em_estom:*

17: gb_gss:*

18: em_gss_hum:*

19: em_gss_inv:*

20: em_gss_pln:*

21: em_gss_vrt:*

22: em_gss_fun:*

23: em_gss_nam:*

24: em_gss_mus:*

25: em_gss_other:*

26: em_gss_pro:*

27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	197.2	13.2	762	14	BQ752115
2	196	13.2	707	14	BQ751801
3	191	12.8	639	14	BQ751512
4	185.2	12.4	742	14	BQ751881
5	183.2	12.3	606	14	BQ751387
6	176.8	11.9	729	10	AW179968

ALIGNMENTS		762 bp mRNA linear EST 18-JUL-2002	
RESULT 1		EST632678 DSCT Colletotrichum trifolii cDNA clone pDSCT10-12, mRNA	
BO752115		BO752115	
LOCUS		LOCUS	
DEFINITION		DEFINITION	
ACCSSION		ACCSSION	
VERSION		VERSION	
KEYWORDS		KEYWORDS	
SOURCE		SOURCE	
ORGANISM		ORGANISM	
REFERENCE		REFERENCE	
AUTHORS		AUTHORS	
TITLE		TITLE	
JOURNAL		JOURNAL	
COMMENT		COMMENT	
FEATURES		FEATURES	
source		source	
1..762	/organism="Colletotrichum trifolii"	1..762	/organism="Colletotrichum trifolii"
/strain="race 1"	/db_xref="taxon:5466"	/strain="race 1"	/db_xref="taxon:5466"
/clone="pDSCT10-12"	/clone_lib="DSCT"	/clone="pDSCT10-12"	/clone_lib="DSCT"
/tissue_type="mycelia"	/dev_stage="young, actively growing mycelia (3 days after inoculation) grown in liquid culture (cutin minimal medium containing 2% glucose)."	/tissue_type="mycelia"	/dev_stage="young, actively growing mycelia (3 days after inoculation) grown in liquid culture (cutin minimal medium containing 2% glucose)."
/lab_host="DH5alpha"	/note="Vector: pBluescript SK+; Site1: EcoRI; Site2: EcoRI; Isolate: 28p2; cDNA was ligated into Lambda gill from enriched RNA The cDNA was ligated into Lambda gill from StrataGene and packaged using Gigapack packaging extracts. An aliquot of the amplified library was used to transduce E. coli Y1090 and phage DNA was purified from a liquid lysate. The cDNA inserts were gel purified from a liquid digestion and ligated into pBluescript SK+. Aliquots of the ligation were used to transform E. coli DH5alpha which were plated onto medium with X-gal for selection of recombinants."	/lab_host="DH5alpha"	/note="Vector: pBluescript SK+; Site1: EcoRI; Site2: EcoRI; Isolate: 28p2; cDNA was ligated into Lambda gill from enriched RNA The cDNA was ligated into Lambda gill from StrataGene and packaged using Gigapack packaging extracts. An aliquot of the amplified library was used to transduce E. coli Y1090 and phage DNA was purified from a liquid digestion and ligated into pBluescript SK+. Aliquots of the ligation were used to transform E. coli DH5alpha which were plated onto medium with X-gal for selection of recombinants."
163 a	267 g	111 t	
BASE COUNT	163 a	267 g	111 t
ORIGIN			
Query Match	13.2%	Score 197.2;	DB 14; Length 762;
Best Local Similarity	62.2%;	Pred. No. 4.3e-51;	
Matches 310;	Conservative 0;	Mismatches 188;	Indels 0; Gaps 0;
QY	632	CTACATATCCAGAACACTGGATTTAACTCCCTCTGAGATTACACAGCCCTGAATCAATCC	691
DB	152	CTGTGGCGGATGGCTGAAGCTCAACATCTCCGCCAGCAGCACGCCGCTCACCACCC	211
QY	692	CTGGGGGCTGATTTGATTATGTCACACAGATTTCAACAGCTGATGAGGCTCTGAA	751
DB	212	GCTGGGCGAGGACTTTGACTACGTGGCGGCTTCAAGTCGCTCGACTACGAGGCGTCAA	271
QY	752	AAAAGATATCAAGATTGCTGACAACTTCCACAGATTGTCCTGGCTGGGATTTGTC	811
DB	272	GAGGACCTACGCCCTGATGACCGACTCCAGGACTGGTGGCTGCGGCTTTGGCCA	331

[illegible]

RESULT 2	707 bp	mRNA	linear	EST '18-JUL-2002
BQ751801				
LOCUS				
DEFINITION	BQ751801	DSCT Colletotrichum trifolii	cdna clone	pDSC7-68, mRNA sequence.

ACCESSION	B0751801
VERSION	B0751801.1
KEYWORDS	GI:21907206
SOURCE	EST.
ORGANISM	Colletotrichum trifolii. Colletotrichum trifolii. Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes; Sordariomycetes incertae sedis; Phyllachorales; Phyllachoraceae; mitosporic Phyllachoraceae; Colletotrichum.

REFERENCE AUTHORS	mitosporic phyriacnaceae; Colletotrichum. 1 (bases 1 to 707) Samac,D.A., Dickman,M., Town,C.D., Van Aken,S., Utterback,T., Cheung,F. and Fraser,C.M.
TITLE	ESTs from mycelia of Colletotrichum trifolii race 1
JOURNAL	Unpublished (2002)
COMMENT	Contact: Deborah A. Samac Department of Plant Pathology University of Minnesota 495 Borlaug Hall, 1991 Upper Buford Circle, St. Paul, MN 55108, USA Tel: 612 625 1243 Fax: 651 649 5058

FEATURES
source 1. .707
Location/Qualifiers
Seq primer: (gTA AtA CgA CtC Act ALA ggg C).
www.medicago.org
TIGR sequence name: MTSAG68TV More information is available at:
Email: debbys@puccini.crl.uiumn.edu
Fax: 631 643 3030

```

FEATURES
source
1. 707
/organism="Colletotrichum trifolii"
/strain="race 1"
/db_xref="taxon:5466"
/clone="pDSC7-68"
/clone_lib="pSC7"
/tissue_type="mycelia"
/dev_stage="Young, actively growing mycelia (3 days after
inoculation) grown in liquid culture (cutin minimal medium
containing 2%glucose)".
/lab_host="DH5alpha"
/notes="Vector: pBluescript SK+; Site.1: EcoRI; Site.2:
EcoRI; isolate: 2sp2 : cDNA was prepared from polyA+
enriched RNA The cDNA was ligated into Lambda gt10 from
Stratagene and packaged using Gigapack packaging extracts.
An aliquot of the amplified library was used to transduce
E. coli Y1090 and phage DNA was purified from a liquid
lysate. The cDNA inserts were gel purified after EcoRI

```

digestion and ligated into pBluescript SK+. Aliquots of the ligation were used to transform *E. coli* DH5alpha which were plated onto medium with X-gal for selection of recombinants.*

RECORDS.			
BASE COUNT	146 a	247 c	195 g
ORIGIN	146 a	247 c	119 t

Query Match 13.2%; Score 196; DB 14; Length 707;

Matches 310; **Conservative** 0; **Mismatches** 190; **Indels** 0; **Gaps** 0;

632 CTACTATCCAGAAACACTGGATTTAACCTCTCTGAGATTACACAGCCCTGAAATCAAATCC 691

178 CTGGTGGCCGGATGCGCTGAAGCTCAACATCTCCGCCAGCACACGCCCGTCACCAACCC 237

QY 692 CTGGGGGCTGATTTTGATTATGCCACCAGATTTCAACAGCTGGGATATGGAGGCTCTGAA 751

Db 238 GCTGGGCCAGGACTTTGACTACGTGCGCGCCTTCAAGTCGCTCGACTACGAGGGCGTCAA 297

752 AAAAGATATCAAAGATTGCTGACAACTTCCCAGGATTGGTGCCCTGCCGATTATGGTCA 811

Db 298 GAAGGACCTCACGGCCCTGATGACCGACTCCCAGGACTGGTGGCTGCCGACTTTGGCCA 357

812 TTATGGTCCTTTCTTTATTTCGTATGGCTTGCGACGGTGCCGGAACATACAGGACATATGA 871

Db
358 CTACGGCGGTCCTGTTTCATCCGCATGGCCCTGGCACAGCGCGGCACGTACCGAGTTCACGA 417

QY 872 TGGCCGGGAGGCCAGTGGTGCTCAGCAACGTTTGTGAACCGCTGAACAGCTGGCCGGA 931

Db 418 CGGACGGGAGGTGGTGGAGAGGGCCAGCAACGGCTTCGCACCGCTCAACAGCTGGCCGGA 477

QY 932 TAACGTTAATCTGGATAAAGCCCGTCGATTGCTGTGGCCAGTCAAGAAAAAATACGGCTC 991

Db 478 CAATGTCAGCCTCGACAAGGCCCGTCGGCTGCTGTGGCCCATCAAGCAAAAGTACGGCAA 537

QY 992 CAGTATTTCCTGGGGAGACCTGATGGTCTGACTGGTAATGTTGCCCTTGAATCCATGGG 1051

Db
538 CAAGATCTCGTGGGCCGACCTCATGATCCTGGCCGGCAACGTGGCCCTCGAGTCCATGGG 597

QY 1052 ATTTAAACGCTGGGATTTCGTCGGCGGAAGAGAAGATGACTGGGAGTCGGACCTGGTATA 1111

Db 598 GTTCCAGACGGCCGGCTTCTCCGGAAGCCGTCCTCCGACACCTGGGAAGCCGAGTCCGT 657

QY 1112 CTGGGGGCTGACACAAGC 1131

Db 658 CTACTGGGGCGCGAGAAC 677

RESULT 3					
BQ751512					
LOCUS	BQ751512	639 bp	mRNA	linear	EST 18-JUL-2002
DEFINITION	EST632075 DSC ^T Colletotrichum trifolii cDNA clone pDSC75-66, mRNA sequence.				

ACCESSION BQ751512
VERSION BQ751512.1 GI:21906917
KEYWORDS EST.
SOURCE Colletotrichum trifolii.
OSCANSISM Colletotrichum trifolii.

ORGANISM Colletotrichum trifolii
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
Sordariomycetes incertae sedis; Phyllachorales; Phyllachoraceae;
mitosporic Phyllachoraceae; Colletotrichum.
REFERENCE 1 (bases 1 to 639)

AUTHORS
Samac, D.A., Dickman, M., Town, C.D., Van Aken, S., Utterback, T.,
Cheung, F. and Fraser, C.M.

TITLE ESTs from mycelia of *Colletotrichum trifolii* race 1
JOURNAL Unpublished (2002)

COMMENT	DATE	BY
Other_ESTs: EST632074		
Contact: Deborah A. Samac		

Contact: Deborah A. Samac
Department of Plant Pathology
University of Minnesota

University of Minnesota
495 Borlaug Hall, 1991 Upper Buford Circle, St. Paul, MN 55108, USA
Tel: 612 625 1243

Fax: 651 649 5058

Email: debbyspuccini.crl.umn.edu
 TIGR sequence name: MTSAE66TV More information is available at:
 www.medicago.org
 Seq primer: (gta Ata Cga Ctc Act Ata ggg C).
 Location/Qualifiers
 1..639
 /organism="Colletotrichum trifolii"
 /strain="race 1"
 /db_xref="taxon:5466"
 /clone="pDST5-66"
 /clone_lib="DST"
 /tissue_type="mycelia"
 /dev_stage="young, actively growing mycelia (3 days after inoculation) grown in liquid culture (cutin minimal medium containing 2% glucose)."
 /lab_host="DH5alpha"
 /note="Vector: pBluescript SK+; Site_1: EcoRI; Site_2: EcoRI; isolate: 2sp2; cDNA was prepared from polyA+ enriched RNA The cDNA was ligated into Lambda gtl from Stratagene and packaged using Gigapack packaging extracts. An aliquot of the amplified library was used to transduce E. coli Y1090 and phage DNA was purified from a liquid lysate. The cDNA inserts were gel purified after EcoRI digestion and ligated into pBluescript SK+. Aliquots of the ligation were used to transform E. coli DH5alpha which were plated onto medium with X-gal for selection of recombinants."
 BASE COUNT 131 a 222 c 176 g 110 t
 ORIGIN

Query Match 12.8%; Score 191; DB 14; Length 639;
 Best Local Similarity 64.2%; Pred. No. 3.8e-49;
 Matches 287; Conservative 0; Mismatches 160; Indels 0; Gaps 0;
 QY 632 CTACTATCAGAAACACTGGATTAACTCTCTGAGATTACACAGCCCTGAATCAATCC 691
 DB 182 CTGGTGGCGGATCGCTGAGCTACATCTCTCCGAGCAGCAGCCGTCACCAACC 241
 QY 692 CTGGGGGCGCTGATTGATTATGATCCACAGATTTCAACAGCTGGATATGAGGCTCTGAA 751
 DB 242 GCTGGCCGAGACTTGTACTAGTCGCGGCTTCAAGTCGCTCAGCTACGAGGCGCTCAA 301
 QY 752 AAAAGATATCAAGATTGCTGACAACTTCCAGAGTTGTGCGCTCGGATATGCTCA 811
 DB 302 GAAGACCTCAGCCCTGATGACCGACTCCAGAGCTGTGGCTCCGACTTGGCCA 361
 QY 812 TTATGGTCTTTCTTTATTCGTATGCTTGGCAGGTGCGGACATACAGGACATGA 871
 DB 362 CTACGGCGGTCTGTTCATCCGCATGGCTGSCACAGCGCGGACGTACCGAGTTCAGA 421
 QY 872 TGGCGGGAGCGCCAGTGTGTGACAGCGTTTGAACCGCTGAACAGCTGGCGGGA 931
 DB 422 CGGACGGGGAGTGGTGAGAGGCGCCAGCAACGCTTCGCCGCTCAACAGCTGGCGGGA 481
 QY 932 TAACGTTAATCTGGATAAGCCCTGCGATGCTGTGGCCAGTCAAGAAAATACGGCTC 991
 DB 482 CAATGTAGCTCAGCAGGCGCTCGGCTGTGTGGCCCATCAAGAAAATACGGCAA 541
 QY 992 CAGTATTTCTGGGAGACGTGATGCTGCTGCTGATGTTAATGTCCTTGAATCCATGGG 1051
 DB 542 CAAGATCTCGTGGCGGACCTCATGATCTTGGCGGCAAGTGCGCTTCAGTCCATGGG 601
 QY 1052 ATTTAAACGCTGGGATTTGCTGGCGG 1078
 DB 602 TTCCAGACGGGCGGCTTCTCCGGAGG 628

RESULT 4
 B0751881
 LOCUS B0751881 742 bp mRNA linear EST 18-JUL-2002
 DEFINITION DSCT Colletotrichum trifolii cDNA clone pDST8-27, mRNA sequence.
 accession B0751881

VERSION
 KEYWORDS
 SOURCE
 ORGANISM

B0751881.1 GI:21907286
 EST
 Colletotrichum trifolii.
 Colletotrichum trifolii
 Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 Sordariomycetes incertae sedis; Phyllachorales; Phyllachoraceae;
 mitosporic Phyllachoraceae; Colletotrichum.
 1 (bases 1 to 742)
 Samac,D.A., Dickman,M., Town,C.D., Van Aken,S., Utterback,T.,
 Cheung,F. and Fraser,C.M.
 ESTs from mycelia of Colletotrichum trifolii race 1
 Unpublished (2002)
 Other ESTs: EST632443
 Contact: Deborah A. Samac
 Department of Plant Pathology
 University of Minnesota
 495 Borlaug Hall, 1991 Upper Buford Circle, St. Paul, MN 55108, USA
 Tel: 612 625 1243
 Fax: 651 649 5058
 Email: debbyspuccini.crl.umn.edu
 TIGR sequence name: MTSAH27TV More information is available at:
 www.medicago.org
 Seq primer: (gta Ata Cga Ctc Act Ata ggg C).
 Location/Qualifiers
 1..742

FEATURES
 source

/organism="Colletotrichum trifolii"
 /strain="race 1"
 /db_xref="taxon:5466"
 /clone="pDST8-27"
 /clone_lib="DST"
 /tissue_type="mycelia"
 /dev_stage="young, actively growing mycelia (3 days after inoculation) grown in liquid culture (cutin minimal medium containing 2% glucose)."
 /lab_host="DH5alpha"
 /note="Vector: pBluescript SK+; Site_1: EcoRI; Site_2: EcoRI; isolate: 2sp2; cDNA was prepared from polyA+ enriched RNA The cDNA was ligated into Lambda gtl from Stratagene and packaged using Gigapack packaging extracts. An aliquot of the amplified library was used to transduce E. coli Y1090 and phage DNA was purified from a liquid lysate. The cDNA inserts were gel purified after EcoRI digestion and ligated into pBluescript SK+. Aliquots of the ligation were used to transform E. coli DH5alpha which were plated onto medium with X-gal for selection of recombinants."
 BASE COUNT 151 a 259 c 212 g 120 t
 ORIGIN

Query Match 12.4%; Score 185.2; DB 14; Length 742;
 Best Local Similarity 62.0%; Pred. No. 2.9e-47;
 Matches 310; Conservative 0; Mismatches 188; Indels 2; Gaps 1;
 QY 632 CTACTATCAGAAACACTGGATTAACTCTCTGAGATTACACAGCCCTGAATCAATCC 691
 DB 178 CTGGTGGCGGATGGCTGAAGCTCAACATCTCCGCCAGCAGCCGCTCACCACCC 237
 QY 692 CTGGGGGCGCTGATTGATTATGCCACAGATTTCACAGCTGGATATGGAGGCTCTGAA 751
 DB 238 GCTGGCCAGAGCTTGTACTAGTCGCCGCTTCAAGTCGCTCGACTACGAGGCGGTCAA 297
 QY 752 AAAAGATATCAAGATTGCTGACAACTTCCAGGATTGGTCCCTGGCGGATTTATGGTCA 811
 DB 298 GAAGGACCTCAGCGGCTGATGACCGACTCCCGAGGACTGGTGGCTCGCGACTTGGCCA 357
 QY 812 TTATGGTCTTTCTTTATTCGTATGCTTGGCAGCGGTGGCGGAAATACAGGACATGA 871
 DB 358 CTACGGCGGTCTGTTCATCCGCATGGCTGGCAGCAGCGCCGACGTACCGAGTTCACGA 417
 QY 872 TGGCGCGGGA--GGCGCCAGTGTGTGTCAGCAACGCTTTTGAACCGCTGAACAGCTGGCGG 929
 DB 418 CGGACGGGAGGGGTGGGAGAGGGCGAGCAACGCTTCGCCACCCCTCAACAGCTGGCGG 477

RESULT 7
 BG370112
 LOCUS
 DEFINITION
 5µ123 V8 mycelial library *Leptosphaeria maculans* cDNA clone Lma123
 Lma123 similar to putative catalase/peroxidase, mRNA sequence.
 ACCESSION
 BG370112
 VERSION
 BG370112.1
 KEYWORDS
 GI:13259728
 SOURCE
 EST.
 ORGANISM
 blackleg of rapeseed fungus.
Leptosphaeria maculans
 Eukaryota; Fungi; Ascomycota; Pezizomycotina; Dothideomycetes;
 Pleosporales; Leptosphaeriaceae; Leptosphaeria.
 1 (bases 1 to 562)
 Iduurm,A., Cozijnsen,A.J. and Howlett,B.J.
 Expressed sequence tags of *Leptosphaeria maculans*, cause of
 blackleg disease of oilseed Brassicas
 Unpublished (2001)
 JOURNAL
 COMMENT
 Contact: Alexander Iduurm
 Barbara J. Howlett (bhowlett@unimelb.edu.au)
 School of Botany
 The University of Melbourne, Parkville, VIC, 3010, Australia
 Tel: 61 3 8344 5056
 Fax: 61 3 9347 5460
 Email: a.iduurm@botany.unimelb.edu.au
 Insert length: 2700 Std Error: 0.00
 Seq primer: T3 Universal Primer.
 Location/Qualifiers
 1..562
 /organism="Leptosphaeria maculans"

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT

Sordariales; Sordariaceae; Neurospora.
 1 (bases 1 to 465)
 Kupfer, D., Lai, H., Nelson, M. and Roe, B.
 ESTs from a Neurospora crassa Sexual cDNA Library
 Unpublished (2001)
 Other ESTs: a3903np.f1
 Contact: Bruce A. Roe, University of Oklahoma, broe@ou.edu
 Department of Chemistry and Biochemistry
 Advanced Center for Genome Technology, University of Oklahoma
 620 Parrington Oval, Norman, OK 73019, USA
 Tel: 405 325 4912
 Fax: 405 325 7762
 Email: broe@ou.edu
 Contact Dr. Mary Anne Nelson, Department of Biology, University of
 New Mexico, Albuquerque, NM 87131 (e-mail address manelson@unm.edu)
 regarding clone availability
 Seq primer: M13 Universal Reverse Primer
 High quality sequence stop: 456.

FEATURES
 source

1. 465
 /organism="Neurospora crassa"
 /strain="wild type"
 /db_xref="taxon:5141"
 /clone="a3903np"
 /clone_lib="Neurospora crassa sexual cDNA library, Uni-zap
 vector system"
 /tissue_type="perithecia (fruiting bodies)"
 /dev_stage="sexual"
 /lab_host="E. coli strain SOLR"
 /note="vector: pBluescript SK-; Site_1: EcoRI; Site_2:
 XhoI; 5' end of cDNA cloned into EcoRI site of pBluescript
 ; 3' end of cDNA cloned into XhoI site of pBluescript"
 93 a 142 c 134 g 96 t

Query Match 10.7%; Score 159.4; DB 12; Length 465;
 Best Local Similarity 69.9%; Pred. No. 4e-39;
 Matches 230; Conservative 0; Mismatches 96; Indels 3; Gaps 1;

Qy 803 TTATGGTCATATGGTCTTTCTTTATTCATGGCTTGGCAGGGTCCGGAACATACAG 862
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 1 TTTGGCCACTATGGCGGTCTCTTCATCCGATGGCATGACAGCGCGGACCTACCG 60

Qy 863 GACATATGTCGGCGGAGCGCCAGTGTGTGTCAGCAACGTTTGAACCGCTGAACAG 922
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 61 TGTACGATGGCGGTGTGTGAGGGTCTATCAGCGTTTGTCTCTCTCAACAG 120

Qy 923 CTGGCCGATAACCTTAATCTGATAAGCCGCTCGATTGCTGTGGCCAGTCAAGAAAA 982
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 121 CTGGCCGACACGCTCTCTCGACAGCGCGTGTCTCTCTGTGGCCATCAAGCAAA 180

Qy 983 ATAGGGTCCAGTATTTCTGGGAGACCTGTGTTGCTGCTGACATGGTAATGTTGCCCTTGA 1042
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 181 GTACGGCAACAAGATCTCGTGGTCCGACCTGCTCTCTTAACCGTTAACCTTGGCTCGA 240

Qy 1043 ATCCATGGATTTAAACGCTGGATTGTCGCGGAGAGAGATGATGAGAGTCGGA 1102
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 241 GTCCATGGGTCTAAGACCTTTGGCTTGGCGGTGGCGGCTCCGACACCTTGGGAGCTGA 300

Qy 1103 CCTG---GTATCTATGGGGCTCTGACACA 1128
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 301 CGAGTCTGTATCTAGGGGTCTGAGACCA 329

RESULT 11
 BQ142683
 LOCUS
 DEFINITION
 accession
 BQ142683
 VERSION
 BQ142683.1
 KEYWORDS
 EST.
 SOURCE
 Metarhizium anisopliae var. acridum.

BQ142683 546 bp mRNA linear EST 24-APR-2002
 113 Metarhizium anisopliae sf. acridum ARSEF 324 Metarhizium
 anisopliae var. acridum cDNA, mRNA sequence.

ORGANISM
 Metarhizium anisopliae var. acridum
 Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 Hypocreales; Clavicipitaceae; mitosporic Clavicipitaceae;
 Metarhizium

REFERENCE
 1 (bases 1 to 546)
 Freilmoser, F.M., Screen, S., Baga, S., Hu, G. and St. Leger, R.J.
 EST analysis of genes expressed by two different insect pathogenic
 fungi during optimized secretion of proteins
 Unpublished (2002)
 Contact: Freilmoser F. M.
 Department of Entomology,
 University of Maryland
 4112 Plant Sciences Building, College Park, MD 20742, USA
 Tel: 301 405 16 13
 Fax: 301 314 92 90
 Email: ff34@umd.edu

FEATURES
 source

1. 546
 /organism="Metarhizium anisopliae var. acridum"
 /strain="ARSEF 324"
 /db_xref="taxon:92637"
 /clone_lib="Metarhizium anisopliae sf. acridum ARSEF 324"
 /note="Vector: UniZap; Metarhizium anisopliae sf. acridum
 was grown on insect cuticle and chitin for 24 hours. A
 cDNA library was constructed in the unidirectional Lambda
 vector UniZap."
 112 a 169 c 149 g 103 t 13 others

Query Match 10.4%; Score 154.6; DB 14; Length 546;
 Best Local Similarity 60.3%; Pred. No. 1.5e-37;
 Matches 297; Conservative 0; Mismatches 169; Indels 27; Gaps 2;

Qy 961 TGCTGTGGCAGTCAGAAAAATACGGCTCCAGTATTTCTGGGGAGACCTGATGCTC 1020
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 4 TGTATGGCCATCAAGCAAAAGTATGGCAAGATCTCTGGCTGTATCTCTGATCC 63

Qy 1021 TGACTGGTAAATGTGGCTTGAATCCATCGGATTTAAACGCTGGGATTTGCTGGCGAA 1080
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 64 TCGCGGCAACGTTGCCCTCGAGTCCATCGGTTTCAAGACATACGGCTTCGCTGGAGGC 123

Qy 1081 GAGAAATGACTGGGAGTCGGACCTGGTATCTATGGGGGCTGACACAAGCC----- 1132
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 124 GTGTTGATGATGGGAGCGGACGAGTCCGCTCTTCTGGGTGGCGAGACACCTGGCTG 183

Qy 1133 -----TCTTGACATTAACGGGATAAAGCAAGCAAGGCTGATCTCGATCCCCC 1176
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 184 GCAACGATGTCCGTTTACAACGACCAACAGGACGTCAGAGCGTGTCTCGATCCCCC 243

Qy 1177 TTGCGCCGACGAGATGGGACCTTATTTATGTCATCTCTGAAGGCCCGGTTGGAACACAG 1236
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 244 TGGCCGCTCCACATGGGTCTCATTTATGTCAACCCGAGGCTCCGATGGAACCCCG 303

Qy 1237 ATCTCTGTGCTCCGGAAGATATCAGGAAGCTTTTTCACGTATGGCCATGATGATG 1296
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 304 ACCCTGTGCTGTCTCGGACATCCGACATCTTGGCCGATGGCCGATGAACGACG 363

Qy 1297 AGGAGACTGTGGCCCTGTATCGGGAGGCGATACATTTGGTAAAGCACATGTCAGACGT 1356
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 364 AGGAGACTGTGGCCCTTATTTGCTGGAGGCCACACCTTNGCAGACATCTACGTTGCCNGN 423

Qy 1357 CTCTCGAAAAATGATTTGGCGGAGG---CCTGATGGTGCACCTGTGGAGGACGAGGAC 1413
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 424 CCTTGNACCAACGTTGGCAAGNAGGCCCCNATGGTGTCCCCCATTTGANCAGAGGCTC 483

Qy 1414 TGGGATGGAAGA 1426
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 484 TNGCTTGAANA 496

RESULT 12
 BH403595/c
 LOCUS
 BH403595 786 bp DNA linear GSS 11-DEC-2001

DEFINITION AG-ND-127D17.TF ND-TAM Anopheles gambiae genomic clone AG-ND-127D17
DNA sequence.
ACCESSION BH403595
VERSION BH403595.1 GI:17349811
KEYWORDS GSS.
SOURCE African malaria mosquito.
ORGANISM Anopheles gambiae
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Anopheles.
REFERENCE 1 (bases 1 to 786)
Shetty,J., Malek,J., Koo,H., Collins,F., Gardner,M. and Loftus,B.J.
Direct Submission of BAC-end sequences from Anopheles gambiae
Unpublished (2001)
Other_GSSs: AG-ND-127D17.TR
Contact: Brendan J Loftus
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0208
Fax: 301 838 3543
Email: bjlloftus@tigr.org
This clone is from an A. gambiae BAC library (ND-TAM) provided by F.H. Collins and sequenced by The Institute for Genomic Research (TIGR). The BAC library was generated from A. gambiae PEST strain DNA. All DNA was extracted from newly hatched first instar larvae to minimize the inclusion of DNA from microorganisms that inhabit the gut. The DNA is derived from mixed sexes of larvae. The BAC library was constructed at Texas A&M University BAC Center University, College Station, Texas 77843-2123, USA using a HindIII partial digest.
Seq primer: M13 For
Class: BAC ends.

FEATURES
source
location/Qualifiers
1..786
/organism="Anopheles gambiae"
/strain="PEST"
/db_xref="taxon:7165"
/clone="AG-ND-127D17"
/clone_lib="ND-TAM"
/note="Vector: pECBAC1; Site_1: HindIII"
BASE COUNT 183 a 211 c 148 g 244 t
ORIGIN

Query Match 10.28; Score 152.6; DB 17; Length 786;
Best Local Similarity 62.09; Pred. No. 8.2e-37;
Matches 298; Conservative 0; Mismatches 169; Indels 14; Gaps 3;
QY 1019 CCTGACTGGTAATGTTGCCCTTGAATCCATGGGATTTAAACGC-TGGGATTTGCTGGCG 1077
DB 729 CCTAACCGGTAACGTTGCTTGGAACTATGGGATTAACAACCAAGGTTATTTGGCGGTG 670
QY 1078 GAAGAGAAGATGACTGGGACTCGGACCTGGTATACTGGGGCGCTGACAAAGCCTCTTG 1137
DB 669 CCGTACAGATGATGAAGACCGATAAGATGTATCTTGGGAGATGAGAAACCTGG 610
QY 1138 CAGATAACCGGGATAAAA-----ACGGGAAACTTCAGAAACCTTTCGCCGCACG 1187
DB 609 CTGGGAACAGATAAAGATATACAGGGGATCGTGATCTGGAAGATCGGCTGGCAGCAACA 550
QY 1188 CAGATGGGACTTATTATGTCATCTGAGAGCCCGCGTGGAAACACAGATCTCTGGCT 1247
DB 549 ACAATGGGGTTTATCTATGTAATCTCTGAAGAGCCGGAAGAAACCGGATCGGCTGGCA 490
QY 1248 TCCGCGAAGATATCAGGGAAGCTTTTTCACGTATGCCATGGATGATGAGGACATGTG 1307
DB 489 GCTGCTAAGATATCTCGTGATACATTTGGCCCGTATGGGATGAATGACGAGGAACCTGTA 430
QY 1308 GCCCTGATCGCGGAGGGCCATACATTTTGGTAAAGACACATGGTGCAGGCTCTCTGAAAAA 1367
DB 429 GCGCTAATTCGTGGTGGCCACACATTTGGTAAACTCATCTGGTGGCGGAGATGACGCCCTA 370
QY 1368 TGTATTGGCAGGGCCTGATGTGTCACCTGTGTGAGAGAGAGGGACTTGGGATGGAATAAT 1427

DB 369 ---GTAGGAGCTGAGCCTGAAGGATTCGGAATTAAGCTCAGGATTAAGTTGGCGAAGT 313
QY 1428 AAATGTGTACAGGAACCGCAATATACCATCACCAGTGCCTTGGAGGAGCCTGCTGTCG 1487
DB 312 AAATTTGACAGAGCTCCGGAGCAGATGCTATTACCAAGTGGTCTGGGAAGTTACCTGGACC 253
QY 1488 A 1488
DB 252 A 252
RESULT 13
AW180185
LOCUS
DEFINITION MGA0268f Mga Library Mycosphaerella graminicola cDNA clone MGA0268
5, similar to PEROXIDASE/CAVITASE, mRNA sequence.
ACCESSION AW180185
VERSION AW180185.1 GI:6447380
KEYWORDS EST.
SOURCE Mycosphaerella graminicola.
ORGANISM Mycosphaerella graminicola
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Dothideomycetes et
Chaetothyrionomycetes incertae sedis; Mycosphaerellaceae;
Mycosphaerella.
REFERENCE 1 (bases 1 to 571)
AUTHORS Keon,J.P.R., Bailey,A.M. and Hargreaves,J.A.
TITLE A group of expressed cDNA sequences from the wheat fungal leaf
blotch pathogen, Mycosphaerella graminicola (Septoria tritici)
JOURNAL Fungal Genet. Biol. 29 (2), 118-133 (2000)
MEDLINE 20374020
COMMENT Contact: Hargreaves JA
Cell Biology Department
IACR-Long Ashton Research Station
Long Ashton, Bristol, BS41 9AF, UK
Tel: +44(0)1275 392181
Fax: +44(0)1275 394281
Email: john.hargreaves@bbsrc.ac.uk
Insert Length: 2500 Std Error: 0.00
Seq primer: M13 reverse.
FEATURES
source
location/Qualifiers
1..571
/organism="Mycosphaerella graminicola"
/strain="Strit"
/db_xref="taxon:54734"
/clone="MGA0268"
/clone_lib="Mga Library"
/note="Vector: pSPORT1; Library constructed from cultures
utilizing ammonium ions as a source of nitrogen"
BASE COUNT 127 a 177 c 141 g 126 t
ORIGIN
Query Match 9.08; Score 134.2; DB 10; Length 571;
Best Local Similarity 61.08; Pred. No. 5.2e-31;
Matches 268; Conservative 0; Mismatches 168; Indels 3; Gaps 3;
QY 632 CTACTATCCAGAAACACTGGATTTAACTCTCTGAGATTACACAGCCCTGAATCAAATCC 691
DB 131 CTGGTGGCCAAATGAGCTGACACCAAGATCTTGGCCACACACCCCTGCTACCGACC 190
QY 692 CTGGGGGGCTGATTTTGTATGATCCACAGATTTCACACAGC-TGGATATGAGGCTCTGA 750
DB 191 ATTCGGCAACGAGTTCGACTACCCAGCGCATCAAGAAGCTTGGATTACAATGGCCCTCA 250
QY 751 AAAAGATATCAAGATTTGCTGACAACTTTCCACAGGATTTGGTCCCTCGCGGATATGGTC 810
DB 251 AGAAGGACTTGAACGACCTCATGACCGACTCGAAGGACTTCTGGCCACGACAGACTTTGGCC 310
QY 811 ATTATGTCCTCTTTTATTTTCGTATGCTTGGCAGCGTTCGGGACATACAGGACATATG 870
DB 311 ACTACGGTGGTTTCTTCGTCCGATGGCATGGCAGCGTGTGATCCCTGCTGCTGCTG 370
QY 871 ATGGCCGGGAGGCGCGAGTGGTGGTCAGCAACCTTTTGAACCGCTGAACAGCTGGCGCG 930

rpm) at room temperature. Culture was harvested, blended, inoculated into minimal medium as above for 24 h. Cells were harvested, washed with water and inoculated into minimal medium base lacking nitrogen source for 6 h. Sequences were processed by phredprap 991019 and trimmed according to phd files and for vector seqs."

89 a 129 c 76 t
BASE COUNT

Db	271	GATCCGGTGTCTGCCGACACGGGATATCCGCACCTTTTAGCCGATGGCCATGAACGAC	330
Qy	1296	GAGGAGACTGTGGCCCTGATCGCGGGAGGCGATACATTGGTAAAGACACATGGTGACGC	1355
Db	331	GAGGAGACCGTCGCTCTGATTGGAGCGGTCACACTGTGGGCAAGACTCACGGTGCTGCA	390
Qy	1356	TCCTCTGAAAA	1366
Db	391	CCCTCTGACAA	401

Dd		391 CCTCTGACAA 401	
RESULT 17			
Bt750328			
LOCUS	Bt750328	718 bp	mRNA linear EST 25-SEP-2001
DEFINITION	Fg02_07d09_R Fg02_AAFc_ECORC_Fusarium-graminearum_mycelium		
	Gibberella zeae cDNA clone Fg02_07d09, mRNA sequence.		
ACCESSION	Bt750328		
VERSION ..	Bt750328.1	GI:15772130	
KEYWORDS	EST.		
SOURCE	Gibberella zeae.		
ORGANISM	Gibberella zeae		

```

FEATURES
    source
        1. 718
            /organism="Gibberella zeae"
            /strain="DAOM 180378"
            /db xref="taxon:5518"
            /clone="Fg02_07d09"
            /clone_lib="Fg02_AAPC_ECORC_Fusarium_graminearum_mycelium"
            /tissue_type="Mycelial tissue"
            /dev stage="Assexual"

```

/lab_host="E. coli (Sure cells)"
 /note="Vector: Bluescript SK+/XhoI-EcoRI; Site_1: EcoRI;
 Site_2: XhoI; Mycelial tissue was collected from V8 agar
 plates after a growth period of 6-7 days at 25 C with 14
 hrs (PL/UV) day light exposure. Mycelia was ground in
 liquid nitrogen prior it's storage at -80 C until RNA
 extraction. Directional cloning with 5' end of cDNA cloned
 into EcoRI site of pBluescript and 3' end of cDNA cloned
 into XhoI site of pBluescript (Stratagene, La Jolla, CA)."
 BASE COUNT 163 a 250 c 185 g 116 t 4 others
 ORIGIN

Query Match 5.9%; Score 87.6; DB 13; Length 718;
 Best Local Similarity 59.5%; Pred. No. 4.8e-16;
 Matches 200; Conservative 1; Mismatches 130; Indels 5; Gaps 3;
 QY 1153 AAAACGGGAACCTTCAGAAACCTCTTCGCCGACGAGATGAGCTATTATGTCAATC 1212
 DB 123 ACAACCGAGATCTCGAAGACCTTTGGCTCTGCTGCCACACGCTCTCATCTATGTTAACC 182
 QY 1213 CTGAAGGCCCGGCTGGAACACAGATCCTCTGCTTCGCGAAGATATCAGGAAGCTT 1272
 DB 183 CTGAGGGCCCTGATGGCAACCCGACCCGTCGCCGCCGCCATGACATCCGCGAGACCT 242
 QY 1273 TTTTCACTGATGGCATGATGAGAGACTGTGGCCCTGATCGCGGGAGGCGCATACAT 1332
 DB 243 TCGGCCCATGG-CATGAACGATGAAGAGACCGTGTGCCCTGATCGCTGTGGTGGCCACAC-G 300
 QY 1333 TTGCTAAGACATGTTGCAGCGTCTCTTGAAATATGTTGGCGAGGCGCTGATGGTG 1392
 DB 301 TCGCAAGACCCAGGAGCTGGCTCAACCGACCA---CGTCGCCGCCGAGCCGAGGCCG 357
 QY 1393 CACCTGTGGAGGACGAGGACTGGGATGGAATAATATGTTACAGGAACGGCAAAAT 1452
 DB 358 CCGACCTTGCCAGCAGGGTCTCGGTGTGTCACAGACTACAAGATTGCAAGGGCCCGC 417
 QY 1453 ATACCATCACCAGTGGCTGGAAGGAGCCCTGGTCTGA 1488
 DB 418 ACAACACACCTCTGGTATCGAAGTCAATGGAACAA 453

RESULT 18
 BM868464
 LOCUS 291 bp mRNA linear EST 07-MAR-2002
 DEFINITION mgc014xp15f.b Magnaporthe grisea CS Uni-Zap XR Library Magnaporthe
 grisea cDNA clone mgc014xp15 5', mRNA sequence.
 ACCESSION BM868464
 VERSION BM868464.1 GI:19236146
 KEYWORDS EST.
 SOURCE Magnaporthe grisea.
 ORGANISM Magnaporthe grisea
 Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 Sordariomycetes incertae sedis; Magnaportheaceae; Magnaporthe.
 REFERENCE 1 (bases 1 to 291)
 AUTHORS Ebbole,D.J., Yuan,J., Thomas,T.L., Bobrowicz,P. and Dean,R.A.
 TITLE Expressed sequence tags from the rice blast fungus, Magnaporthe
 grisea
 JOURNAL Unpublished (2002)
 COMMENT Contact: Ebbole DJ
 Department of Plant Pathology & Microbiology
 Texas A&M University
 Peterson Bldg, MS2132, College Station, TX 77843-2132, USA
 Tel: 979 845 4831
 Fax: 979 845 6483
 Email: d-ebbole@tamu.edu
 Chromatogram file of this sequence is available, see contact person
 ;Best nr hit (Nov. 11, 2001) dbj|BAB03310.1| (AB011415) catalase
 [Sphingomonas sp.] 122 le-27
 PCR Primers
 FORWARD: T3 primer
 BACKWARD: T7 primer
 Plate: mgc014 row: P column: 15
 Seq primer: T3.

FEATURES

Location/Qualifiers
 1. .291
 /organism="Magnaporthe grisea"
 /strain="Gy11"
 /db_xref="taxon:148305"
 /clone="mgc014xp15"
 /clone_lib="Magnaporthe grisea CS Uni-zap XR Library"
 /sex="Mat1-2 hermaphrodite"
 /cell_type="conidia"
 /note="Vector: pBluescriptSK-; Site_1: EcoRI; Site_2: XhoI
 ; Unidirectional cloning. EcoRI side has T3 primer and
 predominantly 5' reads. T7 primer on XhoI side of insert.
 Conidial library. Point inoculation of Gy11 at center of
 oatmeal agar plate. Conidia were harvested after two weeks
 of growth. Sequences were processed by phredphrap 951019
 and trimmed according to phd files and for vector seqs."
 BASE COUNT 63 a 88 c 93 g 47 t
 ORIGIN
 Query Match 5.6%; Score 82.8; DB 14; Length 291;
 Best Local Similarity 64.7%; Pred. No. 9.5e-15;
 Matches 123; Conservative 0; Mismatches 67; Indels 0; Gaps 0;
 QY 948 AAAGCCCGCTGATGCTGTGGCCAGTCAAGAAAAATACGGTCCAGTATTTCCGGGGA 1007
 DB 36 AAAGCTAGGCGGCTGATTTGGCCCATCAAGCAAAAGTATGGCAACAAGATCTCATGGGCC 95
 QY 1008 GACCTGATGCTCTGACTGTAATGTTGCCCTGATCCATCGGATTTAAACGCTGGGA 1067
 DB 96 GACCTGATGCTCTTACGGGCAACGTCGCGCTCGAAGACATGGGCTTCAAGACGCTTGA 155
 QY 1068 TTTGCTGGCGGAAGAAGATGACTGGAGTCGGACCTGATATCTGCGGCGCTGCACAAC 1127
 DB 156 TTGCGGGTGGCAGGCGCCGACAGTGGCAGTCGGAGGCGGCTCTACTGGGTCGCCAG 215
 QY 1128 AAGCCTCTTG 1137
 DB 216 ACCACCTTG 225

RESULT 19

BM862783
 LOCUS 278 bp mRNA linear EST 07-MAR-2002
 DEFINITION mgc004xm19f.b Magnaporthe grisea CM Uni-Zap XR Library Magnaporthe
 grisea cDNA clone mgc004xm19 5', mRNA sequence.
 ACCESSION BM862783
 VERSION BM862783.1 GI:19230465
 KEYWORDS EST.
 SOURCE Magnaporthe grisea.
 ORGANISM Magnaporthe grisea
 Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 Sordariomycetes incertae sedis; Magnaportheaceae; Magnaporthe.
 REFERENCE 1 (bases 1 to 278)
 AUTHORS Ebbole,D.J., Yuan,J., Thomas,T.L., Bobrowicz,P. and Dean,R.A.
 TITLE Expressed sequence tags from the rice blast fungus, Magnaporthe
 grisea
 JOURNAL Unpublished (2002)
 COMMENT Contact: Ebbole DJ
 Department of Plant Pathology & Microbiology
 Texas A&M University
 Peterson Bldg, MS2132, College Station, TX 77843-2132, USA
 Tel: 979 845 4831
 Fax: 979 845 6483
 Email: d-ebbole@tamu.edu
 Chromatogram file of this sequence is available, see contact person
 ;Best nr hit (Nov. 11, 2001) pir|F82584 catalase/peroxidase xf2232
 [imported] - Xylella fast. . . 112 le-24
 PCR Primers
 FORWARD: T3 primer
 BACKWARD: T7 primer
 Plate: mgc004 row: M column: 19
 Seq primer: T3.
 Location/Qualifiers

FEATURES


```

QY 742 AGGCTCTGAAAAAGATATCAAGATTGCTGTGACAACTTCCAGAGATTGTCGCCGCGG 801
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 288 ACGGCTCCGAAAGATATTGTTGACATGCTACCGACTCGAAAGACTACTGCGCTGCTG 347
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 802 ATTATGTCATTATGGCTCTCTTTATTCGTATGGCTTGGCAGCGTGGCGGAACATACA 861
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 348 ACTTGGTAACACGCACCATCATGATCGCTGGCGTGGCAGTGGCGGCTAGTTACC 407
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 862 GGAC--ATATGATGATCGCGGAGCGCCAGTGGTGTGTCAGCAA--CGTTTGAACGCGTG 917
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 408 GCTCTGTTGCGATGCTGAGTGGCTGCGAGGTGGATAGGATCGCTTACGCCGGAG 467
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 918 AACAGCTGGCGGTAACGTTATCTGTGATAAGCCGCTGATGCTGTG 967
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 468 CACGGTTGGCCGGAATAATGCAACCTTGACAAGCGCTTCACATTTTCTG 517
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 22
AW792137 396 bp mRNA linear EST 01-MAY-2001
LOCUS D01025-R Lambda Zap, Stratagene Blumeria graminis f. sp. hordei
DEFINITION cDNA clone D01025 similar to catalase-peroxidase, mRNA sequence.
ACCESSION AW792137
VERSION AW792137.1 GI:13903734
KEYWORDS EST.
SOURCE Blumeria graminis f. sp. hordei.
ORGANISM Eukaryota; Fungi; Ascomycota; Pezizomycotina; Leotiomycetes;
REFERENCE Erysiphales; Erysiphaceae; Blumeria.
AUTHORS Thomas,S.W., Rasmussen,S.W., Glaring,M.A., Rouster,J.A. and Oliver
,R.P.
TITLE Gene identification in the fungal pathogen Blumeria graminis by
expressed sequence tag analysis
JOURNAL Unpublished (2000)
COMMENT Contact: Rasmussen,S.W.
Department of Yeast Genetics
Carlsberg Laboratory
10 Gl. Carlsbergvej, DK-2500, Copenhagen, Denmark
Tel: 45 3327 5230
Fax: 45 3327 4766
Email: swr@erc.dk
High quality sequence stop: 396
POLYA-No.

FEATURES
    source          Location/Qualifiers
     1..396
    /organism="Blumeria graminis f. sp. hordei"
    /db_xref="taxon:62688"
    /clone="D01025"
    /clone_lib="Lambda Zap, Stratagene"
    /cell_type="conidia"
    /lab_host="Hordeum vulgare"
    /lab_host="Hordeum vulgare"
BASE COUNT 118 a 96 c 97 g 85 t
ORIGIN
Query Match 3.7%; Score 55.8; DB 10; Length 396;
Best Local Similarity 59.9%; Pred. No. 5e-06;
Matches 112; Conservative 0; Mismatches 72; Indels 3; Gaps 1;

QY 1302 ACTGTGCCCTGATCGCGGAGGCGCATACATTTGGTAAGCACATGGTCAGGCTCTCT 1361
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1 ACAGTGGCTCTCATTCGCGGTGGACACACTTTTCGCAAGACTCATGGTCCG---CTCCC 57
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1362 GAAAAATGTATTGGCGCAGGCGCTGATGTGTGCACCTGTGGAGGAGCAGGACTGGGATGG 1421
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 58 GCTACTCATCTCGAAGGAACTTGGAGGAGTCCCAATAGACACAGGCTTGGGTGG 117
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1422 AAAAAATAATGTGGTACAGGAACCGCAATATACCATCACAGTGGCTTGAAGGAGCC 1481
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 118 GCTAACACAGCTATCGCTCTGTGAAAAGGTCCGGACACCATACAGTGGTCTTGAAGTCA 177
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1482 TGGTCGA 1488
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 178 TGGACTA 184
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 24
BF739529 399 bp mRNA linear EST 10-JAN-2001
LOCUS NCMI0B4t3 Mycelial Neurospora crassa cDNA clone NM1084 5' similar
DEFINITION to catalase/peroxidase, Streptomyces reticuli, mRNA sequence.
ACCESSION BF739529
VERSION BF739529.1 GI:12066125
KEYWORDS EST.
SOURCE Neurospora crassa.
ORGANISM Neurospora crassa

```

```

Db 178 TGGACTA 184
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 23
AW792143 397 bp mRNA linear EST 01-MAY-2001
LOCUS D01029-R Lambda Zap, Stratagene Blumeria graminis f. sp. hordei
DEFINITION cDNA clone D01029 similar to catalase-peroxidase, mRNA sequence.
ACCESSION AW792143
VERSION AW792143.1 GI:13903740
KEYWORDS EST.
SOURCE Blumeria graminis f. sp. hordei.
ORGANISM Eukaryota; Fungi; Ascomycota; Pezizomycotina; Leotiomycetes;
REFERENCE Erysiphales; Erysiphaceae; Blumeria.
AUTHORS Thomas,S.W., Rasmussen,S.W., Glaring,M.A., Rouster,J.A. and Oliver
,R.P.
TITLE Gene identification in the fungal pathogen Blumeria graminis by
expressed sequence tag analysis
JOURNAL Unpublished (2000)
COMMENT Contact: Rasmussen,S.W.
Department of Yeast Genetics
Carlsberg Laboratory
10 Gl. Carlsbergvej, DK-2500, Copenhagen, Denmark
Tel: 45 3327 5230
Fax: 45 3327 4766
Email: swr@erc.dk
High quality sequence stop: 397
POLYA-No.

FEATURES
    source          Location/Qualifiers
     1..397
    /organism="Blumeria graminis f. sp. hordei"
    /db_xref="taxon:62688"
    /clone="D01029"
    /clone_lib="Lambda Zap, Stratagene"
    /cell_type="conidia"
    /lab_host="Hordeum vulgare"
    /lab_host="Hordeum vulgare"
BASE COUNT 119 a 97 c 96 g 85 t
ORIGIN
Query Match 3.7%; Score 55.8; DB 10; Length 397;
Best Local Similarity 59.9%; Pred. No. 5e-06;
Matches 112; Conservative 0; Mismatches 72; Indels 3; Gaps 1;

QY 1302 ACTGTGCCCTGATCGCGGAGGCGCATACATTTGGTAAGCACATGGTCAGGCTCTCT 1361
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1 ACAGTGGCTCTCATTCGCGGTGGACACACTTTTCGCAAGACTCATGGTCCG---CTCCC 57
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1362 GAAAAATGTATTGGCGCAGGCGCTGATGTGTGCACCTGTGGAGGAGCAGGACTGGGATGG 1421
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 58 GCTACTCATCTCGAAGGAACTTGGAGGAGTCCCAATAGACACAGGCTTGGGTGG 117
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1422 AAAAAATAATGTGGTACAGGAACCGCAATATACCATCACAGTGGCTTGAAGGAGCC 1481
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 118 GCTAACACAGCTATCGCTCTGTGAAAAGGTCCGGACACCATACAGTGGTCTTGAAGTCA 177
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1482 TGGTCGA 1488
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 178 TGGACTA 184
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 24
BF739529 399 bp mRNA linear EST 10-JAN-2001
LOCUS NCMI0B4t3 Mycelial Neurospora crassa cDNA clone NM1084 5' similar
DEFINITION to catalase/peroxidase, Streptomyces reticuli, mRNA sequence.
ACCESSION BF739529
VERSION BF739529.1 GI:12066125
KEYWORDS EST.
SOURCE Neurospora crassa.
ORGANISM Neurospora crassa

```


/note=Vector: pBluescript SK(-); Site_1: EcoRI; Site_2: XhoI

```

BASE COUNT      121 a   145 c   141 g   130 t
ORIGIN
Query Match      3.1%; Score 45.8; DB 12; Length 537;
Best Local Similarity 48.3%; Pred. No. 0.0095;
Matches 128; Conservative 0; Mismatches 137; Indels 0; Gaps 0;

QY 738 ATGGAGGCTCTGAAAAAGATATCAAGATTGCTGACAACTTCCAGGATGGTCCCT 797
Db 110 AAGGAGGACTTCAGAGCTTTATGATGAATGCAAGACTTCTTGTAGAGAGGACGAT 169

QY 798 GCGGATATGTCATTATGCTCTTCTTTATTCGTTATGCTATGGCTGGCAGCGTGC 857
Db 170 TATGATGATGGAAGCTACGAGCACTGCTTGTCTGCTGCTGCGTGGCAGCCAGTGT 229

QY 858 TACAGACATATGATGCGCGGGGAGCGGCGAGTGGTGGTGCAGCAAGTTTGAACGGCTG 917
Db 230 TAGGATAAGGAACCTGGGACTGGAGGAGCAATGGCGCTACCATGGCTTCGCCCATAG 289

QY 918 ACAGCTGGCGGATACGTTAATCTGATTAAGCCGCTCGATGCTGTGGCCAGTCAAG 977
Db 290 TCTGACCAAGGTCCTCAATGCTGATGTAAGGCTGCTAGAGATTCTCTCGAGCCAGTAA 349

QY 978 AAAAATACGGCTCCAGTATTCCT 1002
Db 350 AAGAAATCCCTTGGATTCTTACT 374

```

```

RESULT 29
LOCUS      BI718013
DEFINITION Lambda Zap II Chlamydomonas reinhardtii CC-1690, Stress II (normalized),
ACCESSION BI718013
VERSION    BI718013.1 GI:15693708
KEYWORDS   EST.
SOURCE     Chlamydomonas reinhardtii.
ORGANISM   Chlamydomonas reinhardtii.
REFERENCE  Grossman, A., Chang, C.-W., Davies, J., Harris, E., Hauser, C., Lefebvre
AUTHORS    P., McDermott, J.P., Shrager, J., Silflow, C. and Stern, D.
TITLE      Analyses of the Chlamydomonas reinhardtii Genome: A Model,
            Unicellular System for Analyzing Gene Function and Regulation in
            Vascular Plants. Project: 1031
JOURNAL    Unpublished (2001)
COMMENT     Contact: Charles Hauser
            DCMB Box 91000
            Duke University
            Durham, NC 27708-1000
            Tel: 919 613 8159
            Fax: 919 613 8177
            Email: chauser@duke.edu.

```

```

FEATURES
source
1..660
/organism="Chlamydomonas reinhardtii"
/strain="CC-1690 wild type mt+ 21gr"
/db_xref="taxon:3055"
/clone_lib="C. reinhardtii CC-1690, Stress II (normalized)
), Lambda Zap II"
/note=Vector: pBluescript II SK-; Site_1: EcoRI; Site_2:
XhoI; Stress condition II library, constructed by John
Davies and Jeffrey McDermott, combines cDNAs from CC-1690
cells grown to mid-log phase in TAP (NH4+ - containing)
and shifted to TAP - NO3- (24hrs); H2 production
conditions (0, 12hr, 24hr) see Melis et al., (2000) Plant
Phys. 122: 127-135; TAP + H2O2 (1, 12, 24 hr); TAP +
sorbitol (1, 2, 6, 24 hr); TAP + Cd (1, 2, 6, 24 hr).
PolyA mRNA was purified from each sample, pooled and cDNA
synthesized. The cDNA was directionally cloned into lambda

```

```

Zap II (Stratagene) in the EcoRI (5') and XhoI (3')
sites. pBluescript II SK- plasmids were excised from the
lambda Zap clones by superinfection with ExAssist
(Stratagene) phage. The library was normalized using
method 4 described in Bonaldo et al., (1996) Genome
Research 6: 791-806."
Research 6: 791-806."
BASE COUNT      92 a   231 c   229 g   107 t   1 others
ORIGIN
Query Match      3.0%; Score 44.8; DB 13; Length 660;
Best Local Similarity 53.1%; Pred. No. 0.023;
Matches 94; Conservative 0; Mismatches 83; Indels 0; Gaps 0;

QY 801 GATTATGGTCATTATATGCTCTTCTTTATTCGTTATGCTATGGCTGGCAGCGTGC 860
Db 465 GAGCAGCGGAGCTACGCGCCCGCTGCTGGTGGCTGGCTGGCAGCGCTCCGCGACCTAC 524

QY 861 AGACATATATGTCGGGGGAGCGGCGAGTGGTGGTGCAGCAAGTTTGAACCGCTGAAC 920
Db 525 GCCAAGAGGAGCGGAGCGGCGAGCAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 584

QY 921 AGCTGCGCGGATTAACGTTAATCTGATTAAGCCGCTCGATGCTGTGGCCAGTCAAG 977
Db 585 GAGTGGGCGCCCAACGCGGCTCTGGCGGTCTGCGAAGCTGCTGGAGCCGCTCAAG 641

```

```

RESULT 30
LOCUS      AW789645
DEFINITION C01253-F Lambda Zap, Stratagene Blumeria graminis f. sp. hordei
ACCESSION AW789645
VERSION    AW789645.1 GI:13901242
KEYWORDS   EST.
SOURCE     Blumeria graminis f. sp. hordei.
ORGANISM   Blumeria graminis f. sp. hordei.
REFERENCE  Eukaryota; Fungi; Ascomycota; Pezizomycotina; Leotiomyces;
AUTHORS    Erysiphales; Erysiphaceae; Blumeria.
            Thomas, S.W., Rasmussen, S.W., Glaring, M.A., Rouster, J.A. and Oliver
            , R.P.
TITLE      Gene identification in the fungal pathogen Blumeria graminis by
            expressed sequence tag analysis
JOURNAL    Unpublished (2000)
COMMENT     Contact: Rasmussen, S.W.
            Department of Yeast Genetics
            Carlsberg Laboratory
            10 GL. Carlsbergvej, DK-2500, Copenhagen, Denmark
            Tel: 45 3327 5230
            Fax: 45 3327 4766
            Email: swrecr.dk
            High quality sequence stop: 335
POLYA-No.

```

```

FEATURES
source
1..335
/organism="Blumeria graminis f. sp. hordei"
/db_xref="taxon:62688"
/clone_lib="C01253"
/cell_type="conidia"
/lab_host="Hordeum vulgare"
BASE COUNT      94 a   70 c   78 g   91 t   2 others
ORIGIN
Query Match      2.9%; Score 43.8; DB 10; Length 335;
Best Local Similarity 66.3%; Pred. No. 0.031;
Matches 63; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

QY 1008 GACCTGATGCTCGTACTGTTGTTGCTTGAATCCATGCTTAAACCGCTGGGA 1067
Db 4 GATCTTTACCTATTAACTGTTGTTGCTTGAATCCATGCTTAAACCGCTTGGGA 63

QY 1068 TTGTCTGGCGGAGAGAGATGACTGGGAGTCGGA 1102

```



```

EST.
KEYWORDS      Phaeodactylum tricornutum.
SOURCE        Phaeodactylum tricornutum
ORGANISM      Eukaryota; stramenopiles; Bacillariophyta; Bacillariophyceae;
              Bacillariophycidae; Naviculales; Phaeodactylaceae; Phaeodactylum.

REFERENCE     1 (bases 1 to 694)
AUTHORS       Scala,S., Carels,N., falciatore,A., Chiusano,M.L. and Bowler,C.
TITLE         Genome properties of the diatom Phaeodactylum tricornutum
JOURNAL       Plant Physiol. 129,. (2002) In press
COMMENT       Contact: Bowler C
              Laboratory of Molecular Plant Biology
              Stazione Zoologica 'Anton Dohrn', Italy
              Villa Comunale, I-80121, Napoli, Italy
              Tel: 39 081 583 3268/3211
              Fax: 39 081 764 1355
              Email: chris@alpha.szn.it
              Seq primer: T3 backward.

FEATURES             Location/Qualifiers
source               1..694
                    /organism="Phaeodactylum tricornutum"
                    /db_xref="taxon:2850"
                    /clone_lib="Phaeodactylum tricornutum Uni-Zap XR"
                    /cell_line="CCMP632"
                    /note="Vector: Uni-Zap XR vector; Site_1: Eco RI; Site_2:
                    Xho I"
BASE COUNT          140 a   228 c   192 g   134 t
ORIGIN
Query Match          2.7%; Score 40; DB 14; Length 694;
Best Local Similarity 49.3%; Pred. No. 0.8;
Matches 134; Conservative 0; Mismatches 135; Indels 3; Gaps 1;

QY      825  TTTATTTCGTATGGCTTGGCGGATGCGGGAACATACAGGACATATGATGCCGGGGAGGC 884
          |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db       250  TTTATCCGCTCTCTCGGCATGCTCGGTACCTACAGTAAGCCGACATTCGGCGGT 309

QY      885  GCCAGTGGTGGTCAGCAACGCTTTTGAACCGCTGAACAGCTGGCCGGGATAACGTTAACTG 944
          |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db       310  TCGAACGGTGGCGGTATGCTTTTATCCCGAAGCGGATGGGCGCCAACGCCGGACTC 369

QY      945  GATAAAGCCCGTCGATGCTGTGCCAGTCAAGAAAAANTACGGCTCCAGTATTTCCTGG 1004
          |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db       370  AAAGTAGCCCGCGACGCCCTTGGAAACCGCTCAAGGCCAAGTTC---CCGGTCTATCGTAC 426

QY      1005  GGAGACCTGATGGTCTCTGACTGTGTAATGTTGCCCTTGAATCCATGGGATTTAAACCGTG 1064
          |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db       427  GCGACCTCTACACCTACGCTGGTGTGTCGTGCGGTGGAAGCCGAGGCCCTATCATTT 486

QY      1065  GGATTTGCTGGCGGAAGAGAAGATGACTGGGA 1096

Db       487  CCCTTTGCCACGGCGCCGACGGACGAATCGGA 518

```

RESULT 34	BH119903/c
LOCUS	BH119903
DEFINITION	RPCI-24-316M7.TV RPCI-24 linear GSS 19-JUL-2001
ACCESSION	BH119903
VERSION	RPCI-24-316M7.Mus musculus genomic clone RPCI-24-316M7,
KEYWORDS	DNA sequence.
SOURCE	BH119903 GI:14963415
ORGANISM	GSS. house mouse. Mus musculus Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 543) Zhao,S., Nierman,W., Malek,J., Shatsman,S., Akinret,B., Levins,M., Tsegaye,G., Geer,K., Krol,M., Shivartsbeyn,A., Gebregeorgis,E., Russell,D., de Jong,P. and Fraser,C.M. Mouse BAC End sequences from Library RPCI-24 Unpublished (1999)
TITLE	Other_GSSs: RPCI-24-316M7.TV
JOURNAL	Contact: Shaying Zhao
COMMENT	

Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org
Clones are derived from the mouse BAC library RBCI-24. For BAC library availability, please contact Pieter de Jong (pdj@engmail.cho.org). Clones may be purchased from BACPAC Resources (<http://www.choi.org/bacpac/orderingframe.htm>). BAC end page: http://www.tigr.org/tadb/bac_ends/mouse/bac_end_intro.html
Plate: 316 row: M column: 7
Seq primer: SP6
Class: BAC ends.

FEATURES	source	BASE COUNT	ORIGIN
Class: BAC ends:	Location/Qualifiers	203 a	76 c 94 g 170 t
	1. .543		
	/organism="Mus musculus"		
	/strain="C57BL/6J"		
	/db_xref="taxon:10090"		
	/clone="RPCI-24-316M7"		
	/clone_lib="RPCI-24"		
	/sex="Male"		
	/cell_type="Spleen/Brain"		
	/notes="Vector: pTARBAC1; Site_1: BamH1; Site_2: BamH1; RPCI-24 Mouse BAC library produced by Pieter de Jong. The library was cloned in the pTARBAC1 cloning vector at the BamH1 sites using KhoI partially digested male C57BL/6J DNA."		

Query Match	2.7%	Score 39.8;	DB 17;	Length 543;
Best Local Similarity	50.8%;	Pred. No. 0.8;		
Matches	95;	Conservative 0;	Mismatches 92;	Indels 0; Gaps 0;
QY	590	TTCTACCGCTGTAGCCGCTGATAAAAAGAGAGCTCAAAATTTCTACTATCCACAAACACT	649	
Db	423	TTGTCACACAGAGATCCGATTAACATGAAGATTCAAAATCACTCAACTCAATAATC	364	
QY	650	GGATTAACTCTCTGAGATTACACAGCCCTGAATCAAAATCCTCGGGGCGCTGATTTGA	709	
Db	363	AGGGTCAAAATATTTTGTCTCTCTCTCTCATTAACACAGCTCTTGGTTACATTTTCTGAA	304	
QY	710	TTATGCCACACAGATTTCAACAGCTGGATATGGAGGCTCTGAAAAAGATATCAAGAATTT	769	
Db	303	ATTTTAGACAGAATGTCCCACTGGACATATAGTTTATATATAAAGTTCTCTGAAATTT	244	
QY	770	GCTGACA	776	
Db	243	ATTTTCCA	237	

RESULT 35	
BG711981	
LOCUS	
DEFINITION	pglln.pk010.g23.Normalized Liver Library Gallus gallus cDNA clone
VERSION	pglln.pk010.g23 5' similar to dbj BAB14712.1 (AK023886) unnamed
KEYWORDS	protein product [Homo sapiens]G, mRNA sequence.
SOURCE	EST.
ORGANISM	chicken.
	Gallus gallus
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
	Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
	Phasianinae; Gallus.
	1 (bases 1 to 661)
REFERENCES	Burnside,J., Morgan,R.W. and Cogburn,L.A.
AUTHORS	Chicken ESTs from a normalized liver library
TITLE	Unpublished (2001)
JOURNAL	Contact: Joan Burnside
COMMENT	Molecular Endocrinology
	University of Delaware

```

pBelOBAC11.
FEATURES
  source
    Location/Qualifiers
      1. .803
        /organism="Drosophila melanogaster"
        /db_xref="taxon:7227"
        /clone="BACN13P04"
        /clone_lib="DrosBAC"
        /plasmid="pBelOBAC11"
        /note="end : 17"
BASE COUNT      119 a      113 c      128 g      138 t      305 others
ORIGIN
Query Match      2.6%; Score 39.4; DB 17; Length 803;
Best Local Similarity 19.0%; Pred. No. 1.4;

```


COMMENT
- Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <http://www.fruitfly.org> The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Mammos in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain Y2; cn bw sp, the same strain used for the BDGP's p1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

Public Soybean EST Project
Unpublished (1999)
Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.edu
This clone is available through: ResGen, Invitrogen Corp. 2130 South Memorial Parkway Huntville, AL 35801 For further information call: (800)-533-4363 or contact via email: ccu@resgen.com
Insert length: 1196 Std Error: 0.00
High quality sequence stop: 422.

FEATURES

source
1..1095
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone="BACR22A01"
/clone_lib="RPCI-98"
/note="end : 77"
Location/Qualifiers

BASE COUNT 171 a 66 c 98 g 233 t 527 others
ORIGIN

Query Match 2.6%; Score 38.2; DB 17; Length 1095;
Best Local Similarity 22.4%; Pred. NO. 4;
Matches 57; Conservative 29; Mismatches 169; Indels 0; Gaps 0;

QY 292 TATTTCAGATACCTTCGTATCAACACGTAACACAGAGACACAGCTTTTGTCT 351
DB 441 TTTNTSSSSNNNNNNSSSANNNAANNNSSNNNTSSNNNNNANAANNTNNNT 500
QY 352 GACATCCCAAGAGAGGAGATATTCAGGTCTGCGCAGCACATCAACGATCGTCAGTGC 411
DB 501 NANNNSSSGTNTTNNSSNNNTNNNNNNNNNNNNNNSSSSSAAANNNNSNS 560
QY 412 GCCTGGAAACCCCTAGTATTTTGTGTAGTATCTATCCAGCAATAGGTATATCCT 471
DB 561 SNTTNNNAANNAATTTTSSSTTTTSSSTTTTAAAAAANNNNNNNNNATAAT 620
QY 472 GTTGCATCAAAAGTTCACATTTGTATACACATGCGAATTCCTTAATCCGAGCTA 531
DB 621 TANNTANNTTAAATTTTAAATTTTAAACSSNNNNSSNNNNNNNNNNNNNT 680
QY 532 TTCGTATGATAAAA 546
DB 681 TSSSNTTNTAVAA 695

RESULT 41

AW707196 599 bp mRNA linear EST 03-DEC-2001
LOCUS sk22a06.y1 Gm-c1028 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:
DEFINITION Gm-c1028-2891 5' similar to TR:Q39780 Q39780 ASCORBATE PEROXIDASE.
; mRNA sequence.

ACCESSION

AW707196 1 GI:7591464

VERSION

EST.

KEYWORDS

soybean.

SOURCE

Glycine max

ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;

Glycine.

1 (bases 1 to 599)

REFERENCE

1

Shoemaker, R., Keim, P., Vodkin, L., Erpelting, J., Coryell, V., Khanna

, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C.,

Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers

, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk

, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann

, R., Waterston, R. and Wilson, R.

BASE COUNT

172 a 131 c 146 g 148 t 2 others

ORIGIN

Query Match 2.5%; Score 37.8; DB 10; Length 599;
Best Local Similarity 50.9%; Pred. NO. 3.7;
Matches 143; Conservative 0; Mismatches 132; Indels 6; Gaps 2;

QY

819 CCTTCTTTTATCGTATGCTTGGCAGCGTCCGGACATACAGACATATGATGCCGG 878

DB

135 CCTCTCATGCTTCGATTAGCTGGCATGATGCTGGTACTTACGATGCTAAAAACAACA 194

QY

879 GGAGGCCCGCAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 938

DB

195 GGAGGCCCTTAATGGTTCTATCAGAACACAGACAGAGATTGAATCAC---GCACCAACACAG 251

QY

939 ATCTGGGATAAGCCCGTGCATTCGTGTCGCCAGTCAAGAAAAATACGGTCCAGTATT 998

DB

252 GGACTGGAAACACGACCTTCCTCTCTGTAGGAAGTGAAGGCCCAACA---TCCAAAAAT 308

QY

999 TCTTGGGGAGCCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1058

DB

309 TCATATGCCGACCTTTACCAGCTAGCTGTGTGTTGTCAGTAGAAGTACCAGGGGGGCCA 368

819 CTTTCTTTTATTGCTATGCTTGCACGGTGCCGGAACATACAGACATATGATGCCGG 878
 150 CTCTCTCATGCTTTCGATTACCGTGGCATGCTGGTACTTACGATGCTAAACAACAACA 209

```

Db      122 CCTTGTGTGATCTGTTCCAGCAATACTGAATAAGAGTTGAGCCTGTTGATGCC 181
QY      1390 GTGCACCTGTGGA 1402
Db      182 GCCCAGCTGCAGA 194

RESULT 44
LOCUS   AW180016
DEFINITION MG0072f Mga Library Mycosphaerella graminicola cDNA clone Mga0072
5', similar to catalase-peroxidase, mRNA sequence.
ACCESSION AW180016
VERSION   AW180016.1 GI:6447211
KEYWORDS EST.
SOURCE   Mycosphaerella graminicola.
ORGANISM Mycosphaerella graminicola.
          Eukaryota; Fungi; Ascomycota; Pezizomycotina; Dothideomycetes et
          Chaetochyriomycetes incertae sedis; Mycosphaerellaceae;
          Mycosphaerella.
REFERENCE 1 (bases 1 to 627)
AUTHORS  Keon,J.P.R., Bailey,A.M. and Hargreaves,J.A.
TITLE    A group of expressed cDNA sequences from the wheat fungal leaf
          blotch pathogen, Mycosphaerella graminicola (Septoria tritici)
JOURNAL  Fungal Genet. Biol. 29 (2), 118-133 (2000)
MEDLINE  20374020
COMMENT  Contact: Hargreaves JA
          Cell Biology Department
          IACR-Long Ashton Research Station
          Long Ashton, Bristol, BS41 9AF, UK
          Tel: +44(0)1275 392181
          Fax: +44(0)1275 394281
          Email: john.hargreaves@bsrc.ac.uk
          Insert Length: 2500 Std Error: 0.00
          Seq primer: M13 reverse.

FEATURES             source
    source            1..627
                        /organism="Mycosphaerella graminicola"
                        /strain="Strit"
                        /db_xref="taxon:54734"
                        /clone="MGA0072"
                        /clone_lib="Mga Library"
                        /note="Vector: pSPORT1; Library constructed from cultures
                        utilizing ammonium ions as a source of nitrogen"

BASE COUNT    146 a 204 c 160 g 117 t
ORIGIN

Query Match      2.5%; Score 37.6; DB 10; Length 627;
Best Local Similarity 59.3%; Pred. No. 4.4;
Matches 64; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY      1382 GCCTGATGGTGCACCTGTGGAGGAGCAGGACTGGGATGGAATAAATGTGGTACAGG 1441
Db      1 GCCTAAACGGCGCCCTCCATCGAGGAGGAGGCTCGGATGGAAGACAGTACAGGACGG 60

QY      1442 AAACGGCAATATACATCACCAGTGCCTGGAGGAGCCTGGTCGAC 1489
Db      61 CAAGGGCTCAACGCCCATCATCCGCGATTGGAAGTCATCTGGACGTC 108

RESULT 45
LOCUS   DR22018S/C
DEFINITION Danio rerio genomic clone Dkey-22018, genomic survey sequence.
ACCESSION AL745989
VERSION   AL745989.1 GI:21349287
KEYWORDS GSS.
SOURCE   zebrafish.
ORGANISM Danio rerio
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
          Cypriniformes; Cyprinidae; Danio.
REFERENCE 1 (bases 1 to 630)

```

```

AUTHORS  Humphray,S.J., Huckle,E. and Hunt,S.E.
TITLE    Direct Submission
JOURNAL  Submitted (06-JUN-2002) The Sanger Institute, Wellcome Trust Genome
          Campus, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail contact:
          humquery@sanger.ac.uk Unpublished
COMMENT  This sequence was generated from the SP6 end of BAC 22018. 22018 is
          part of the Daniokey Pilot BAC Library created by R. Plasterk and
          N.V. Keygene.
          Further details: http://www.sanger.ac.uk/Projects/D_rerio/.

FEATURES             source
    source            1..630
                        /organism="Danio rerio"
                        /db_xref="taxon:7955"
                        /clone="Dkey-22018"
                        /tissue_type="Testis"
                        /note="vector pindigobAC-536"

BASE COUNT    259 a 59 c 82 g 230 t
ORIGIN

Query Match      2.5%; Score 37.6; DB 17; Length 630;
Best Local Similarity 47.1%; Pred. No. 4.4;
Matches 115; Conservative 0; Mismatches 129; Indels 0; Gaps 0;

QY      446 ATCTATCCCGACGATAGGTATATATCTGTTGTCATCATATAAGTTGACTTTTGTATACAACA 505
Db      630 ATATATACATGGAGTATGAGTATCATATTTGTTGTCTTACTTTTAAATTCAGACCTTATA 571

QY      506 TCGGAATTTCCCTTAATCCGGAGCTATTCGTATGATAAAAAAACHCTTCCCTGTTCTGAT 565
Db      570 CACTTATATATTTATATACACTTATTCATCATATATTAATTAATGATGCTCATTT 511

QY      566 TCTTCGGCGCTATCCGGGAGCTTTTCTACCGCTGTAGCGCTGATAAAAAAGAGACTCA 625
Db      510 TCTTCTTGACATTCAGATGATCTTTTTTGTTCATTAACACACTCATTAATTTAATAGC 451

QY      626 AAATTTCTACTATCCAGAAACACTGGATTTAACTCTCTGAGATTACACAGCCCTGAATC 685
Db      450 ATGTATCATTTTAACTAATTAACACAGCTTTTACCATTAAATTAATTAACACATCATAA 391

QY      686 AAT 689
Db      390 TAAT 387

RESULT 46
LOCUS   BQ611766
DEFINITION sap65a03.y1 Gm-cl087 Glycine max cDNA clone SOYBEAN CLONE ID:
          Gm-cl087-5382 5' similar to TR:Q9SXL5 Q9SXL5 CHLOROPLAST ASCORBATE
          PEROXIDASE PRECURSOR ;, mRNA sequence.
ACCESSION BQ611766
VERSION   BQ611766.1 GI:21601435
KEYWORDS EST.
SOURCE   soybean.
ORGANISM Glycine max
          Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
          Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
          Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
          Glycine.
REFERENCE 1 (bases 1 to 587)
AUTHORS  Shoemaker,R., Keim,P., Vodkin,L., Erpellding,J., Corvelli,V., Khanna
          ,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,
          Wyllie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers
          ,Y., Person,B., Swallier,T., Gibbons,M., Pape,D., Harvey,N., Schurk
          ,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann
          ,R., Waterston,R. and Wilson,R.
          Public Soybean EST Project
          Unpublished (1999)
CONTACT: Shoemaker R/Public Soybean EST Project
          Public Soybean EST Project
          Washington University School of Medicine
          4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
          Tel: 314 286 1800

```

CDNA Library Arrayed by: The I.M.A.G.E. Consortium/LLNL
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL, send email to:
info@image.llnl.gov

```

Seq primer: -40UP from Gibco
High quality sequence stop: 475.
Location/Qualifiers
1. .476
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3174440"
/clone_lib="NCI CGAP_Brn41"
/tissue_type="oligodendroglioma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: brain; Vector: p7T73P-pac (Pharmacia) w
modified polylinker; Site.1: NotI; Site.2: EcoRI; 1st
strand cDNA was primed with a Not I - oligo(gtT) primer
TGTTCACCATCTGAGTGGCAGCGCGGCATACATTTTTTTTTTTTTTTT
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the
I and Eco RI sites of the modified p7T73 vector. Lib
constructed and normalized by Bento Soares and M.Fat
Bonaldo."

```

BASE COUNT		128 a	124 c	136 g	84 t	4 others	
ORIGIN							
Query Match 2.5%; Score 37; DB 12; Length 476;							
Best Local Similarity 51.2%; Pred. No. 5.8;							
Matches 108; Conservative 0; Mismatches 99; Indels 4; Gaps 1;							
QY	1211	TCCTGAAGGCCCGGTGGAAACACAGATCCTCTGGCTTCGCCGAAAGATATCAGGGAAGC	1270				
DB	253	TGCTCAGGCCCCCATAGGATAACAGATTGGTGGGCCCTGGAGATAGAGGAGAGNAAG	312				
QY	1271	TTTTTTCACGATGGCCATGGATGATGAGGAGACTGTGGCCCTGATCGCGGAGGGCATAC	1330				
DB	313	TATATGTGGTACTGGGAGGAGGAAGGAAGTANCCNNTTAGAGCCACTTTTCCAT	372				
QY	1331	ATTTGGTAAAGCACATGGTCACCGTCTCCTGAAAAA----	1386				
DB	373	CTTTGGTTTCTACAGGCTGGATGGTATCCCTTGGAAAGCCTTGAGCCCCACAGCCCTG	432				
QY	1387	ATGGTCACACTGTGGAGGACAGGGACTGGG	1417				
DB	433	GTAGGGCTCTCTGGAAGAGCAAGGGACAGCG	463				

AW981658	AW981658	520 bp	mRNA	linear	EST 02-JUN-2000
LOCUS	PC15B10	Pine Triplex pollen cone library	Pinus taeda	clone	
DEFINITION	PC15B10, mRNA sequence.				
ACCESSION	AW981658				
VERSION	AW981658.1	GI:8173224			
KEYWORDS	EST.				
SOURCE	loblolly pine.				
ORGANISM	Pinus taeda				
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus. 1 (bases 1 to 520)				
REFERENCE	Whetten, R.W., Kintaw, C.S., Retzel, E. and Sederoff, R.R.				
AUTHORS	The Pine Gene Discovery Project				
TITLE	Unpublished (1999)				
JOURNAL					

CONTACT: Ross Whetten
Forest Biotechnology Group
North Carolina State University
Dept. of Forestry, NC State University, 6113 Jordan Hall, Raleigh
, NC, 27695-8008
Tel: 919-515-7800
Fax: 919-515-7801
Email: rosswhetten@ncsu.edu
Seq primer: 5' lambda Triplex2 Sequencing Primer.

FEATURES

```

location/qualifiers
1. 520
/organism="Pinus taeda"
/db_xref="taxon:3352"
/clone="PC15810"
/clone_lib="Pine Triplex pollen cone library"
/dev_stage="immature"
/lab_host="E. coli BM25.8"
/note="Organ: pollen cone; Vector: Lambda Triplex; Site:1:
SfiI (A); Site_2: SfiI (B); Immature pollen cones were
collected in the early spring, frozen and used for mRNA
isolation. The SMART-PCR method (Clontech) was used to
prepare a library from 1 ug total RNA, using the Lambda
Triplex vector. Plasmid subclones in pTriplex were
recovered by cre-lox excision in E. coli strain BM25.8 and
sequenced from the 5' end."
139 a 115 c 136 g 121 t 9 others
BASE COUNT

```

Query Match	2.5%	Score 37	DB 10	Length 520
Best Match	100%	100%	100%	100%
Next Best Match	95%	95%	95%	95%
Third Best Match	90%	90%	90%	90%
Fourth Best Match	85%	85%	85%	85%
Fifth Best Match	80%	80%	80%	80%
Sixth Best Match	75%	75%	75%	75%
Seventh Best Match	70%	70%	70%	70%
Eighth Best Match	65%	65%	65%	65%
Ninth Best Match	60%	60%	60%	60%
Tenth Best Match	55%	55%	55%	55%
Eleventh Best Match	50%	50%	50%	50%
Twelfth Best Match	45%	45%	45%	45%
Thirteenth Best Match	40%	40%	40%	40%
Fourteenth Best Match	35%	35%	35%	35%
Fifteenth Best Match	30%	30%	30%	30%
Sixteenth Best Match	25%	25%	25%	25%
Seventeenth Best Match	20%	20%	20%	20%
Eighteenth Best Match	15%	15%	15%	15%
Nineteenth Best Match	10%	10%	10%	10%
Twentieth Best Match	5%	5%	5%	5%
Twenty-first Best Match	0%	0%	0%	0%
Twenty-second Best Match	0%	0%	0%	0%
Twenty-third Best Match	0%	0%	0%	0%
Twenty-fourth Best Match	0%	0%	0%	0%
Twenty-fifth Best Match	0%	0%	0%	0%
Twenty-sixth Best Match	0%	0%	0%	0%
Twenty-seventh Best Match	0%	0%	0%	0%
Twenty-eighth Best Match	0%	0%	0%	0%
Twenty-ninth Best Match	0%	0%	0%	0%
Thirtieth Best Match	0%	0%	0%	0%
Thirty-first Best Match	0%	0%	0%	0%
Thirty-second Best Match	0%	0%	0%	0%
Thirty-third Best Match	0%	0%	0%	0%
Thirty-fourth Best Match	0%	0%	0%	0%
Thirty-fifth Best Match	0%	0%	0%	0%
Thirty-sixth Best Match	0%	0%	0%	0%
Thirty-seventh Best Match	0%	0%	0%	0%
Thirty-eighth Best Match	0%	0%	0%	0%
Thirty-ninth Best Match	0%	0%	0%	0%
Fortieth Best Match	0%	0%	0%	0%
Forty-first Best Match	0%	0%	0%	0%
Forty-second Best Match	0%	0%	0%	0%
Forty-third Best Match	0%	0%	0%	0%
Forty-fourth Best Match	0%	0%	0%	0%
Forty-fifth Best Match	0%	0%	0%	0%
Forty-sixth Best Match	0%	0%	0%	0%
Forty-seventh Best Match	0%	0%	0%	0%
Forty-eighth Best Match	0%	0%	0%	0%
Forty-ninth Best Match	0%	0%	0%	0%
Fiftieth Best Match	0%	0%	0%	0%
Fifty-first Best Match	0%	0%	0%	0%
Fifty-second Best Match	0%	0%	0%	0%
Fifty-third Best Match	0%	0%	0%	0%
Fifty-fourth Best Match	0%	0%	0%	0%
Fifty-fifth Best Match	0%	0%	0%	0%
Fifty-sixth Best Match	0%	0%	0%	0%
Fifty-seventh Best Match	0%	0%	0%	0%
Fifty-eighth Best Match	0%	0%	0%	0%
Fifty-ninth Best Match	0%	0%	0%	0%
Sixtieth Best Match	0%	0%	0%	0%
Sixty-first Best Match	0%	0%	0%	0%
Sixty-second Best Match	0%	0%	0%	0%
Sixty-third Best Match	0%	0%	0%	0%
Sixty-fourth Best Match	0%	0%	0%	0%
Sixty-fifth Best Match	0%	0%	0%	0%
Sixty-sixth Best Match	0%	0%	0%	0%
Sixty-seventh Best Match	0%	0%	0%	0%
Sixty-eighth Best Match	0%	0%	0%	0%
Sixty-ninth Best Match	0%	0%	0%	0%
Seventieth Best Match	0%	0%	0%	0%
Seventy-first Best Match	0%	0%	0%	0%
Seventy-second Best Match	0%	0%	0%	0%
Seventy-third Best Match	0%	0%	0%	0%
Seventy-fourth Best Match	0%	0%	0%	0%
Seventy-fifth Best Match	0%	0%	0%	0%
Seventy-sixth Best Match	0%	0%	0%	0%
Seventy-seventh Best Match	0%	0%	0%	0%
Seventy-eighth Best Match	0%	0%	0%	0%
Seventy-ninth Best Match	0%	0%	0%	0%
Eightieth Best Match	0%	0%	0%	0%
Eighty-first Best Match	0%	0%	0%	0%
Eighty-second Best Match	0%	0%	0%	0%
Eighty-third Best Match	0%	0%	0%	0%
Eighty-fourth Best Match	0%	0%	0%	0%
Eighty-fifth Best Match	0%	0%	0%	0%
Eighty-sixth Best Match	0%	0%	0%	0%

[illegible]

RESULT 49	BH762769/c	BH762769	637 bp	DNA	linear	GSS 20-MAR-2002
LOCUS	BH762769	BH762769				
DEFINITION	BMBAC330H08T7_P5U Brugia malayi Genomic Bac Library 3 Brugia malayi genomic, DNA sequence.					
ACCESSION	BH762769					
VERSION	BH762769.1					
KEYWORDS	GSS.					
SOURCE	Brugia malayi.					
ORGANISM	Brugia malayi.					
REFERENCE	Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea; Onchocercidae; Brugia.					
AUTHORS	1 (bases 1 to 637) Whitton,C., Daub,J., Ware,J., Quail,M., Hall,N., Barrell,B., Foster J., Guillemin,D., Siatko,B. and Blaxter,M.					
TITLE	Genome survey sequences from the human parasitic nematode Brugia					

**JOURNAL
COMMENT**

malayi
Unpublished (2000)
Contact: Blaxter ML
Institute of Cell, Animal and Population Biology
University of Edinburgh
Ashworth Labs, King's Buildings, West Mains Road, Edinburgh, EH9
3JT, UK
Tel.: +44 131 650 6760
Fax: +44 131 670 5450
Email: mark.blaxter@ed.ac.uk
Sequenced from the Brugia malayi BAC library constructed by Clai
Whitton and Dr Mike Quail. The sequence was generated by The
Pathogen Sequencing Unit, The Sanger Institute, Cambridge, UK in
collaboration with Mark Blaxter, IACPB, University of Edinburgh,
Edinburgh, UK.
Seq primer: T7 (TAATACGACTACTACTAGGG)

FEATURES	SOURCE
-----------------	---------------

```

Location/Qualifiers
1. .637
/organism="Brugia malayi"
/strain="TRS"
/db_xref="taxon:6279"
/clone_lib="Brugia malayi Genomic Bac Library 3"
/sex="Mixed (male and female)"
/tissue_type="whole parasite"
/dev_stage="microfilaria (L1)"
/note="Vector: pBACE3.6; Site_1: BamH I; Brugia malayi
genomic DNA was partially cleaved with Sau3A I and size
fractionated. 7,392 clones were generated with mean insert
size ~48 kbp. The library was constructed by Claire
Whitton, Blaxter Nematode Genetics Lab, University of
Edinburgh, UK, and Dr Mike Quail, The Pathogen Sequencing
Unit, The Sanger Centre, Cambridge, UK."

```

Query Match 2.5%; Score 37; DB 17; Length 637;
Best Local Similarity 50.9%; Pred. No. 6.9;
Matches 88; Conservative 0; Mismatches 85; Indels

617	AGAGACTCAAAATTTCTACTATCCAGAAACACTGGATTTAACTCTCTCAGATTACACAG	676
324	ACAGACTAAATAATTTTGCACCACTTAACCTTCAATAATAAATGACAAAAATCACAGCA	265
677	CCCTGAATCAAAATCCCTGGGGGCTGATTTTGATTATGCCACCAAGATTTCAACAGCTGGA	736
264	CTATTAGATCATACCTTCAGATCATTTTGAGAAATCAACCACTTGTGTAATAATTC	205
737	TATGGAGGCTCTGAATAAAGATATCAAGATTTGCTGCACAACTTCCAGGATT	789
204	TGTGGATTGTCTCAAAACGACATATAAATCTCGTCGAGCTGCATTCATT	152

RESULT 50
CNC00E.DM

CNS0005BT	1024 bp	DNA	linear	GSS 03-JUN-1999
LOCUS				
DEFINITION	Drosophila melanogaster genome survey sequence T7 end of BAC # BACR11006 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.			
ACCESSION	AL058090			
VERSION	AL058090.1	GI:4932671		
KEYWORDS	GSS.			
SOURCE	Drosophila melanogaster.			
ORGANISM	Drosophila melanogaster.			
	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.			
REFERENCE	1 (bases 1 to 1024)			

REFERENCE
AUTHORS
TITLE
JOURNAL

THIS PAGE BLANK (USPTO)

Db 301 ATACCTTCGTCAACAGCTACAAACAGACAGCAGCTTTTGTTCGTGACATCCAC 360
 QY 361 AAGAGAGGAATATTCAGGTCTCGCAGCACTCAACGGCACTCAGTTCAGTTCGGCTGGAA 420
 Db 361 AAGAGAGGAATATTCAGGTCTCGCAGCACTCAACGGCACTCAGTTCAGTTCGGCTGGAA 420
 QY 421 CCCTTTAGTATTTTCTGTAGTATCTATCCAGCAATAGTATATCTGTGATCA 480
 Db 421 CCCTTTAGTATTTTCTGTAGTATCTATCCAGCAATAGTATATCTGTGATCA 480
 QY 481 ATAAAGTGTACTTTTCTATACAACTGCGAATTTCCCTTAATCCGGAGCTATTTCGTATGA 540
 Db 481 ATAAAGTGTACTTTTCTATACAACTGCGAATTTCCCTTAATCCGGAGCTATTTCGTATGA 540
 QY 541 TAAAAAAACTCTTCCTGTTCTGATTCCTGCGGTATCGGGAGCTTTTCTACCGCTG 600
 Db 541 TAAAAAAACTCTTCCTGTTCTGATTCCTGCGGTATCGGGAGCTTTTCTACCGCTG 600
 QY 601 TAGCCGCTGATAAAAAAGAGACTCAAAATTTCTACTATCCAGAAACACTGGATTAACTC 660
 Db 601 TAGCCGCTGATAAAAAAGAGACTCAAAATTTCTACTATCCAGAAACACTGGATTAACTC 660
 QY 661 CTCTGAGATTACACAGCCCTGAATCAATCCCTGGGGGCTGATTTGATTATGCCACCA 720
 Db 661 CTCTGAGATTACACAGCCCTGAATCAATCCCTGGGGGCTGATTTGATTATGCCACCA 720
 QY 721 GATTTCACAGCTGGATGAGGCTCTGAAAAAGATATCAAAAGATTTGCTGACAACTT 780
 Db 721 GATTTCACAGCTGGATGAGGCTCTGAAAAAGATATCAAAAGATTTGCTGACAACTT 780
 QY 781 CCCAGATTGGTCCCTGCGGATATGTTCAATTTGTTCTTCTTTATTCGTATGGCTT 840
 Db 781 CCCAGATTGGTCCCTGCGGATATGTTCAATTTGTTCTTCTTTATTCGTATGGCTT 840
 QY 841 GGCACGCTGCGGAACATACAGACATATGATGCGCGGGAGCGCCAGTGGTGGCAGC 900
 Db 841 GGCACGCTGCGGAACATACAGACATATGATGCGCGGGAGCGCCAGTGGTGGCAGC 900
 QY 901 AACGTTTGAACCGCTGAACAGCTGCGCGGATTAACGTTTATCTGGATTAAGCCCGTCGAT 960
 Db 901 AACGTTTGAACCGCTGAACAGCTGCGCGGATTAACGTTTATCTGGATTAAGCCCGTCGAT 960
 QY 961 TGTCTGGCCAGTCAAGAAAAATACGGTCCAGTATTTCTGGGAGACCTGATGGTCC 1020
 Db 961 TGTCTGGCCAGTCAAGAAAAATACGGTCCAGTATTTCTGGGAGACCTGATGGTCC 1020
 QY 1021 TGACTGATATGTTGCCCTTGAATCCATGCGATTTAAACGCTGGGATTTGCTGGCGGAA 1080
 Db 1021 TGACTGATATGTTGCCCTTGAATCCATGCGATTTAAACGCTGGGATTTGCTGGCGGAA 1080
 QY 1081 GAGAAGTACTGGGAGTGGACCTGGTATCTGGGGGCTGACAAAGCCCTTTGGCAG 1140
 Db 1081 GAGAAGTACTGGGAGTGGACCTGGTATCTGGGGGCTGACAAAGCCCTTTGGCAG 1140
 QY 1141 ATACCGGGATAAAACGGGAATCTCAGAACTCTTCCGCCACGACAGATGGACTTAA 1200
 Db 1141 ATACCGGGATAAAACGGGAATCTCAGAACTCTTCCGCCACGACAGATGGACTTAA 1200
 QY 1201 TTTATGTCATCTGTAAGGCCCGGTGGAACACAGATCTCTGCGCTTCGCCGAAAGATA 1260
 Db 1201 TTTATGTCATCTGTAAGGCCCGGTGGAACACAGATCTCTGCGCTTCGCCGAAAGATA 1260
 QY 1261 TCAGGGAAGCTTTTTCACGTATGGCCATGGATGATGAGGAGACTGTGGCCCTGATCGCG 1320
 Db 1261 TCAGGGAAGCTTTTTCACGTATGGCCATGGATGATGAGGAGACTGTGGCCCTGATCGCG 1320
 QY 1321 GAGGGATACATTTGTTAAGACATGTCACCGCTCTCTGAAATATGTTATGGCGCAG 1380
 Db 1321 GAGGGATACATTTGTTAAGACATGTCACCGCTCTCTGAAATATGTTATGGCGCAG 1380
 QY 1381 GGCTGATGTTGACCTGTGGAGGAGCAGGAGCTGGGATGGAAAAATAAATGTTGGTACAG 1440
 Db 1381 GGCTGATGTTGACCTGTGGAGGAGCAGGAGCTGGGATGGAAAAATAAATGTTGGTACAG 1440

QY 1441 GAAACGGCAATATACATCACCAGTGGCTGGAAGAGAGCTGGTGCAC 1489
 Db 1441 GAAACGGCAATATACATCACCAGTGGCTGGAAGAGAGCTGGTGCAC 1489

RESULT 2

AAV06555
 ID AAV06555 standard; DNA; 2238 BP.

XX AAV06555;

XX 03-JUL-1998 (first entry)

XX Microscilla furvens catalase-53cal gene.

XX Catalase; epoxidation; hydroxylation; biosensor; paper bleaching; pasteurisation; ss.

XX Microscilla furvens.

XX Key Location/Qualifiers

FT CDS 1..2238

FT /tag- a

FT /product- "Catalase-53cal"

XX WO9800526-A1.

XX PN 08-JAN-1998.

XX 03-JUL-1997; 97WO-US16513.

XX 03-JUL-1996; 96US-0674887.

XX (RECO-) RECOMBINANT BIOCATALYSIS INC.

XX Adhikary RS, Robertson DE, Sanyal I;

XX WPI; 1998-086953/08.

XX P-PSDB; AAW33810.

XX New bacterial catalases, related nucleic acid vectors and transformed cells - used as oxidising agents and for detecting or destroying hydrogen peroxide, e.g. in biosensors

XX Claim 3; Fig 2; 35pp; English.

XX The present sequence is of the Microscilla furvens catalase-53cal gene. Fragments of the gene can be used to identify related sequences. Catalase-53cal may be used to catalyse oxidation reactions such as epoxidation or hydroxylation. The enzyme can also be used to detect or destroy hydrogen peroxide, e.g. in connection with glyoxylic acid production, biosensors, contact lens cleaning, pulp/paper bleaching and pasteurisation of dairy products. Antibodies raised against catalase-53cal can be used to screen libraries for detection and purification of cells containing the enzyme.

XX Sequence 2238 BP; 634 A; 545 C; 605 G; 454 T; 0 other;

Query Match 24.0%; Score 357.6; DB 19; Length 2238;
 Best Local Similarity 66.7%; Pred. No. 2.7e-105;
 Matches 542; Conservative 0; Mismatches 264; Indels 6; Gaps 2;

QY 683 ATCAAAATCCCTGGGGGCTGATTTGATTATGCCACAGATTTCAACAGCTGGATGGA 742
 Db 168 ATCGGACCCAAACAGCCCGGATTTGACTATGCCGAGAGATTTAAGAGCTAGATCTGGC 227
 QY 743 GGCTCTGAAAAAGATATCAAGATTTGCTGACAACCTCCAGAGTTGGTCCCTGCGGA 802
 Db 228 AGCGGTTAAAAAGGACCTGGCAGCGCTAATGACAGATTTACAGGACTGGTGGCCAGCA 287
 QY 803 TTATGTCATATGGTCCCTTTCTTTATTCCTATGGCTTGGCCAGCTGGCCGAACATACAG 862

Db 288 TTACGGTCATTATGCCCCCTCTTTATACGATGGCGTGGCACAGCCCGCACCTACCG 347
 QY 863 GACATATGATGCGCGGAGCGCCAGTGTGGTCAACAAGCTTTTGAACCGCTGAACAG 922
 Db 348 TATCGGTATGCGCGTGGCGGTGGTCCGGTCAACGGCTTCGGCTCTCAATAG 407
 QY 923 CTGGCCGATACCTTATCTGATATAAGCCCGCTCATGTTGCTGGCCAGTCAAGAAAA 982
 Db 408 CTGGCCAGACAATGCCAATCTGGATAAAGCACGCTTCTTCTTGGCCCATCAACAAAA 467
 QY 983 ATACGGCTCCAGTATTTCTGGGAGAGACCTGATGGTCTGACGTGGTAATGTTGCCCTGA 1042
 Db 468 ATACGGTGAATAATCTCTGGCGGATCTAATGATATCTACAGGAACAGTAGCTCTGGA 527
 QY 1043 ATCCATGGGATTTAAACGCTGGGATTTGCTGGCGGAAGAGAACATGACCTGGAGTC-- 1099
 Db 528 AACTATGGGCTTTAAACTTTTGGTTTGGAGTGGCAGACAGATGATGGGAGCTGA 587
 QY 1100 GGACCTGTATACCTGGGCGCTGACAAAGCCTTTGTCAGATTAACCGGATATAA---AA 1156
 Db 588 AGAAGATGTATCTGGGAGCAGAAACCGAATGGCTGGGAGACAAAGCGCTATGAAGTGA 647
 QY 1157 CGGGAACCTTCAGAAACCTTTGCGCCACGACATGGACTTATTTATGTCATCTCTGA 1216
 Db 648 CGAGAGCTCGAAATATCCCTGGGAGCGGTACAAATGGGACTCATCTATGTAACCCCGA 707
 QY 1217 AGGCCCCGGTGGAAACACAGATCTCTGGCTTCCGCGAAGATATACAGGAAGCTTTTTC 1276
 Db 708 AGGACCCACGCGAAGCCAGACCTATGCTGCTGCGGTGATATCTGAGACTTTTGG 767
 QY 1277 AGTATGCGCATGATGATGAGGAGACTGTGGCCCTGATCGCGGAGGGGATACATTGG 1336
 Db 768 CGAATGCAATGATGACGAAGAACCGGTGGCTCTCATAGCGGGTGGACACACCTTCGG 827
 QY 1337 TAAACACATGTCGAGCGCTCTCTGAAATATGTTGGCGCAGGCGCTGATGTGTCACC 1396
 Db 828 AAAACCCATGTTGTCGCGATGCGGAGAAATATGTTGGCGCAGAGCTTCCCGCGGAGG 887
 QY 1397 TGTGGAGGACGAGGACTGGGATGGAAATAAATATGTTGATGACGAAACCGCAATATAC 1456
 Db 888 TATTGAAGAAATGAGCTGGGTGGAAACACCTAGCGCACCGGACGCTGCGGATAC 947
 QY 1457 CATCACCATGTCCTGGAGGAGCTGTGCGA 1488
 Db 948 CATCACCATGTCAGTAGAAGCGCGCTGGACCA 979

RESULT 3

AAT89967
 ID AAT89967 standard; DNA; 2235 BP.
 AC AAT89967;
 XX AAT89967;
 XX 20-MAR-1998 (first entry)
 DT
 DE Mycobacterium tuberculosis partial katG gene.
 XX
 XX Tuberculosis; katG gene; isoniazid resistance; INH;
 KW Isonicotinic acid hydrazide; ss.
 XX
 OS Mycobacterium tuberculosis.
 XX
 XX Key Location/Qualifiers
 PH 85..2235
 FT /*tag= a
 FT /product= "katG gene"
 FT /note= "partial cds"
 XX
 XX US5688639-A.
 XX
 XX 18-NOV-1997.
 XX
 XX 18-APR-1994; 94US-0228662.

XX 18-APR-1994; 94US-0228662.
 PR (MAYO-) MAYO FOUNDATION.
 PA
 XX Cockerill FR, Kline BC, Uhl JR;
 PI WPI; 1998-007975/01.
 XX
 DR Determination of isoniazid sensitivity of Mycobacterium tuberculosis
 PT strains - by restriction length polymorphism analysis of katG gene
 PT
 XX Claim 1; Column 9-12; 18pp; English.
 PS
 XX This partial DNA sequence encodes the katG gene of Mycobacterium
 CC tuberculosis strain H37Rv which is used in a novel method to rapidly
 CC identify strains of M. tuberculosis which are resistant to isoniazid
 CC (INH, isonicotinic acid hydrazide). The method involves the use of
 CC restriction fragment length polymorphism (RFLP) analysis to determine
 CC if a NciI-MspI restriction site is absent in the DNA of the strain at
 CC the codon corresponding to codon 463 of a M. tuberculosis katG gene
 CC consensus sequence (see AAW31343). The absence of the site indicates an
 CC INH-resistant strain.
 XX
 XX Sequence 2235 BP; 431 A; 596 C; 745 G; 363 T; 0 other;

Query Match 20.0%; Score 298.4; DB 19; Length 2235;
 Best Local Similarity 62.4%; Pred. No. 5.5e-86;
 Matches 502; Conservative 0; Mismatches 296; Indels 6; Gaps 2;

QY 688 ATCCCTGGGGGCTGATTTTGTATATGCCACACAGATTTCACAGCTGGATATGGAGCTC 747
 Db 176 ACCGATGGGTGGCGGTTCGACTATGCCGCGAGGTCCGACCATCGAGCTTGGACGCC 235
 QY 748 TAAAAAAGATATCAAGATTTCTGACAACTTCCAGGATTTGTCCTTCCGATATATG 807
 Db 236 TGACGCGGACATCGAGGAGTATGATCCACCTCGCAGCGGTGTGCGCCGCGACTAG 295
 QY 808 GTCAATATGTCCTTCTTTTATGTTATGTTGTCGACGCTGCGGAGATACAGACAT 867
 Db 296 GCCACTACGGCGCGCTGTTTATCCGATGGCGTGGCAGCGTCCGCGCACTACCGCATCC 355
 QY 868 ATGATGCGCGGAGCGCCAGTGGTGGTTCAGCAAGCTTTTGAACCGCTGAACAGCTGC 927
 Db 356 ACGACG 415
 QY 928 CGGATAACCTTAATCTGGATAAAGCCCGTGGATTTGCTGGCCAGTCAAGAAAAATACG 987
 Db 416 CCGACAACGCCAGCTTGGACAAGCGCGCGCGCTGCTGTGCGCGGTCAAGAGAAGTACG 475
 QY 988 GCTCCAGTATTTCTGGGAGACCTGATGTCCTGCTGACTGATGTTGCCCTTGAATCCA 1047
 Db 476 GCAAGAAGCTCTCATGGCGGACCTGATTTTTCGCCGCAACTCGCGCGCTGGAATCGA 535
 QY 1048 TGGGATTTAAACGCTGGGATTTGCTGGCGGAAGAGAAGATGACTGGGATCGGACCTGG 1107
 Db 536 TGGCTTCAAGACGTTCCGGTTCCGGTTCGCGCGGTTCGACCATCGGAGCCCGATGAGG 595
 QY 1108 TATACCTGGGCGCTGACAAACGCTTTCGACATTAACCGGATA---AAACGGGAAC 1164
 Db 596 TCTATTGGGCAAGGAAGCCACCTGGCTCGCGCATGAGCGTTTACAGCGGTAAACGGGATC 655
 QY 1165 TTCAGAAACCTTTCGCGCCACGACATGGAATTTATTTATTTTCAATCTCTGAAGGCCCG 1224
 Db 656 TGGAGAACCCCGTGGCGCGCGTGCAGATGGGCTGATCTACGTGAACCCGAGAGGCCCGA 715
 QY 1225 GTGGAAACACAGATCTCTTGGCTTCCGCGAAGATATACAGGAAGCTTTTTCACGTATGG 1284
 Db 716 ACGCAACCCGACCCCATGCGCGCGGTGACATTCGCGAGACGTTTTCGCGCATG 775
 QY 1285 CCATGATGATGAGGAGACTGTGCCCTGATCGGCGGAGGCGATACATTTGGTAAAGCAC 1344
 Db 776 CCATGAACGAGCTGCAAAACAGCGCGCTGATCGTGGCGGTCTACACTTTCGCGTAAGACC 835

896 ATGGCGCCG---GCCCGCGCGATCTGTGTCGGCCCGGAAACCCGAGGCTGTCCGCTGGAGC 952
 1405 AGCAGGACCTGGGATGGAATAAATGCTGTACAGGAACGCAAAATATACCATCACCA 1464
 953 AGATGGGCTTGGGCTGGAAGAGCTGTATGGCACCGGACCGTAAGGACCGCATCACCA 1012
 1465 GTGGCCTGGAAGGAGCCTGTGTCGA 1488
 1013 GCGGCATCGAGGCTGTATGGACGA 1036

RESULT 5

AAAX00817
 ID AAAX00817 standard; DNA; 2331 BP.

AC AAAX00817;

DT 26-MAR-1999 (first entry)

DE M. tuberculosis catalase peroxide (katG) gene sequence.

KW Catalase-peroxide; katG; isonicotinic acid hydrazide; tuberculosis;

KW INH; diagnosis; detection; ss.

OS Mycobacterium tuberculosis.

PH Key Location/Qualifiers
 CDS 70..2292

FT /*tag= "katG"

FT /gene= "catalase peroxide"

FT /product= "catalase peroxide"

FT /note= "the start codon is not indicated"

XX WO980585-A1.

XX 12-NOV-1998.

XX 06-MAY-1998; 98WO-US09285.

XX 07-MAY-1997; 97US-0852219.

XX (MAYO-) MAYO FOUNDATION.

XX Cockerill FR, Kline BC, Uhl JR;

XX WPI; 1999-070099/06.

XX P-PSDB; AAW95398.

XX Detection of Mycobacterium tuberculosis - by amplifying katG gene

XX and detecting specific fragment, and optionally identifying

XX INH-resistant strains by detecting specific mutation

XX Claim 3; Fig 7; 83pp; English.

XX The invention relates to a novel method of detecting Mycobacterium
 XX tuberculosis. The method comprises amplifying the DNA in the samples to
 XX generate a detectable amount of amplified DNA comprising a catalase-
 XX peroxide (katG) DNA fragment with sequence of bases 904-1523 of the
 XX M. tuberculosis strain H37Rv MC katG gene sequence. The method optionally
 XX further comprises determining if the katG DNA fragment has a serine to
 XX threonine mutation in codon 315 (S315T mutation), indicative of an
 XX isonicotinic acid hydrazide (INH)-resistant M. tuberculosis strain. The
 XX method can be used to detect M. tuberculosis in biological fluids,
 XX especially human sputum, useful to diagnose tuberculosis. This disease is
 XX a major cause of human morbidity and mortality, and conclusive diagnosis
 XX and subsequent treatment depends on identification of the etiologic agent
 XX M. tuberculosis. INH has been used in tuberculosis treatment, but
 XX INH-resistant strains have emerged; the method allows such drug-resistant
 XX strains to be identified. The present sequence represents the wild-type
 XX M. tuberculosis katG gene sequence.

XX Sequence 2331 BP; 455 A; 734 C; 763 G; 379 T; 0 other;

Query Match 20.0%; Score 298.4; DB 20; Length 2331;
 Best Local Similarity 62.4%; Pred. No. 5.6e-86;
 Matches 502; Conservative 0; Mismatches 296; Indels 6; Gaps 2;
 QY 688 ATCCCTGGGGGCTGATTTTGGATATGCCACCAGATTTCAACAGCTGGATATGGAGGCTC 747
 DB 236 ACCGATGGGTGGGCTTCGACTATGCGCGGAGGTGCGGACCATCGACGTTGAGGCC 295
 QY 748 TGAATAAAGATATCAAGATTTTCTGACAACTTCCACAGATTTGGTCCCTGCGGATTATG 807
 DB 296 TGACGCGGCACATCGAGGAAGTGTACACCACTCGCAGCGGTGTGTGCGCCCGGACTAG 355
 QY 808 GTCAATATGCTCTTCTTTATTCGTATGCTGTGGCAGCTGCGGGAACATACAGACAT 867
 DB 356 GCCACTACGGGCGCTGTTTATCCGGATGCGTGGCAGCTGCCGCACTTACCGCATCC 415
 QY 868 ATGATGCGCGGGAGGCGCGAGTGGTGGTCAACAGCTTTTGAACCGCTGAACAGCTGC 927
 DB 416 ACAGCGCGCGCGCGCGCGCGGCGGCGCATGCGGCGGTTCGCGCGCTTACAGCTGC 475
 QY 928 CGGATAACCTTAATCTGATATAAGCCCTCGATTGCTGTGGCCAGTCAAGAAAAATACG 987
 DB 476 CCGACAACCGCCAGCTTGGACAAGGCGCGCGCTGCTGTGGCGGTCAAGAAAGTACG 535
 QY 988 GCTCCAGTATTTCTCGGGAGACCTGATGCTGCTGACTGCTGATGTTGCTCCCTGATCCA 1047
 DB 536 GCAAGAAGCTCTCATGCGCGGACCTGATTTTTCGCCGCGCAACTGCGCGCTGGAATCA 595
 QY 1048 TGGGATTTAAACGCTGGGATTTGCTGGCGGAAGAAGATGACTGGAGTGGAGCTGG 1107
 DB 596 TGGGCTTCAAGAGTTTCGGGTTTCGGCTTCGGCGGTTCGACCACTGGGAGCCGATGAG 655
 QY 1108 TATACTGGGGGCTGACAAAGCCCTCTTGCAGATAAACCGGGATA---AAAAGGGGAAC 1164
 DB 656 TCTATTGGGCAAGGAAGCACCTGGCTCGGCGATGAGCGTTACAGCGGTAAAGCGGATC 715
 QY 1165 TTCAGAAACCTCTTTCGGCCACGACAGTGGACTTATTTATGTCATCTCTGAGGCCCG 1224
 DB 716 TGGAGAACCGCTGGCGCGGCTGATGATGATGATGATGATGATGATGATGATGATGAT 775
 QY 1225 GTGGAACACAGATCTCTGGCTTCGCGGAAAGATATCAGGGAAGCTTTTTCACGTATGG 1284
 DB 776 ACGCAACCGGACCCCATGGCGCGGCTGCACATTCGCGAGACGTTTCGCGCATGG 835
 QY 1285 CCATGATGATGAGGAGATGTCGCCCTGATCGCGGAGGCGATACATTTGTTAAAGCAC 1344
 DB 836 CCATGAACGACGTCGAACACAGCGCGCTGATGCTCGCGGCTCACACTTCGGTAAGACC 895
 QY 1345 ATGGTGCAGCGCTCTCTGMAAAATGTTTGGCGCAGGCGCTGATGCTGACCTGTGGAGG 1404
 DB 896 ATGGCGCG---GCCCGCGGATCTGTCGCCCGCCGACCCGAGGCTGCTCCCGTGAGC 952
 QY 1405 AGCAGGACTGGGATGGAATAAATAATGTTGTTACAGGAACCGCAAAATATACCATCACCA 1464
 DB 953 AGATGGGCTTGGGCTGGAAGAGCTCGTATGGCACCGGAACCGTAAGGACCGCATCACCA 1012
 QY 1465 GTGGCTGGAAGAGCGCTGTGTCGA 1488
 DB 1013 GCGGCATCGAGGCTGTATGGACGA 1036

RESULT 6

AAI99683/C

ID AAI99683 standard; DNA; 4403765 BP.

XX AAI99683;

XX AC AC

XX 15-JAN-2002 (first entry)

XX Mycobacterium tuberculosis strain H37Rv genome SEQ ID NO 2.

XX Mycobacterium tuberculosis; strain H37Rv; strain CDC 1551; genome;

XX variation; epidemiology; patient treatment; epidemic monitoring; ds.

XX

XX

XX

XX

XX

CC monitoring.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from USPTO
CC at seqdata.uspto.gov/sequence.html?DocID=6294328B1.
XX SQ Sequence 4411529 BP; 758565 A; 1449983 C; 1444602 G; 758379 T; 0 other;
Query Match 20.0%; Score 298.4; DB 22; Length 4411529;
Best Local Similarity 62.4%; Pred. No. 6.2e-84;
Matches 502; Conservative 0; Mismatches 296; Indels 6; Gaps 2;
QY 688 ATCCCTGGGGGCTGATTTGATTATGATCCACAGATTTCACAGCTGGATGAGGCTC 747
DB 2155943 ACCCGATGGTGGCGGTTCGATATGCGCGGAGTCCGACCATCGACGTTGACGCC 2155884
QY 748 TGAATAAGATATCAAGATTGCTGACAACTCCAGATTGCTGCCCTCGGATATG 807
DB 2155883 TGACCGGACATCAGGAAGTGTGACCACTCGACCCGTGTGGCCCGCCGACATG 2155824
QY 808 GTCAATTATGGTCTTTCTTTATTCGTATGGCTTGGCAGGTGCGGAAACATACAGGACAT 867
DB 2155823 GCCACTACGGCGCTGTTTATCCGATGGCTGCGACGCTGCGGACCTAACGCATCC 2155764
QY 868 ATGATGGCGGGAGCGCGAGTGTGTCAACACGTTTGAACCGCTGAACAGCTGGC 927
DB 2155763 ACACGCGCGCGCGCGCGCGCGCGCATGCGACGCTGCGCGCGCTTAAACAGCTGGC 2155704
QY 928 CGGATAACGTTAATCTGGATAAGCCGCTGCTGATTCGTCGACAGTCAAGAAAAATACG 987
DB 2155703 CGGACACGCCAGCTTGACAAAGCGCGCGGCTGCTGCGCGGTCAAGAAAGTACG 2155644
QY 988 GTCCAGTATTTCCCTGGGGAGACCTGATGGTCTGACTGGTAAATGTTGCCCTTGAATCCA 1047
DB 2155643 GCAAGAAGCTCTCATGGCGGACCTGATTTGTTTGGCGGCAACTGCGCGCTGGAATCGA 2155584
QY 1048 TCGGATTTAAACGCTGGATTTGCTGGCGAAGAGATGACTGGGAGTCGGACCTGG 1107
DB 2155583 TGGGCTTCAAGACGTTCCGGTTTCGGCTTCGGCGGCTCGACAGTGGGAGCCGATGAGG 2155524
QY 1108 TATCTGGGGGCTGACACACAGCTCTTGACAGATAACCGGATA---AAACGCGGAAC 1164
DB 2155523 TCTATTGGGCAAGAAAGCCACCTGGCTCGCGGATGAGCGTTACAGCGGTAAGCGGATC 2155464
QY 1165 TTCAGAACTCTTGGCGCCAGCAGATGGGACTTATTATGTCATCTCAATCTGAAGGCCCG 1224
DB 2155463 TGGAGAACCGCTGGCGCGGTGCAGATGGGGCTGATCTACGTGAACCCGAGGGGCGGA 2155404
QY 1225 GTGGAAACACAGATCTCTGGCTTCGCGGAAAGATATCAGGAAGCTTTTTCACGTATGG 1284
DB 2155403 ACGGCNACCGACCCCATGGCGCGGCGGTGCACATTCGCGAGACGTTTCGGCGCATGG 2155344
QY 1285 CCATGATGATGAGGAGACTGTGGCCCTGATCGCGGGAGGGGATACATTGGTAAGCAC 1344
DB 2155343 CCATGAACACGCTGAAACAGCGCGCTGATCGTCGGCGGTACACATTTTCGTAAGACCC 2155284
QY 1345 ATGGTGCACGCTCTCTGAAAATGTTATGGCGGAGGCGCTGATGGTCCACCTGTGGAGG 1404
DB 2155283 ATGGGCGG---GCCGCGCGATCTGGTGGCGCGAACCAGGAGCTGCTCGCTGGAGC 2155227
QY 1405 ACAGGCGACTGGATGGAATAAATGTTGTGACAGAAACGGCAATATACCATCACCACCA 1464
DB 2155226 AGATGGCTTGGGTGGAAGACCTCGTATGGCACCGGAACCGGTAAGGACGCGATCACC 2155167
QY 1465 GTGGCTTGAAGAGCCTGGTTCGA 1488
DB 2155166 GGGCATCAGGTCGATGAGCA 2155143
RESULT 8
AA000818
ID AA000818 standard; DNA; 2331 BP.
XX
AC AA000818;

XX DT 29-MAR-1999 (first entry)
XX DE M. tuberculosis catalase peroxide (katG) gene variant.
XX KW Catalase-peroxide; katG; mutation; isonicotinic acid hydrazide;
XX KW INH; tuberculosis; diagnosis; detection; variant; ss.
XX OS Mycobacterium tuberculosis.
XX OS Synthetic.
XX FH Key Location/Qualifiers
XX CDS 70..2292
XX FT /*tag= a
XX FT /product= "catalase peroxide variant"
XX FT /note= "the start codon is not indicated"
XX FT 1013
XX FT variation
XX FT /*tag= b
XX FT /note= "G to C variation as compared to the
XX wild-type katG gene"
XX W09850585-A1.
XX PN 12-NOV-1998.
XX PD 06-MAY-1998; 98WO-US09285.
XX PF 07-MAY-1997; 97US-0852219.
XX PR (MAYO-) MAYO FOUNDATION.
XX PA Cockerill FR, Kline BC, Uhl JR;
XX PI WPI; 1999-070099/06.
XX DR P-PSDB; AAW95399.
XX XX Detection of Mycobacterium tuberculosis - by amplifying katG gene
XX and detecting specific fragment, and optionally identifying
XX INH-resistant strains by detecting specific mutation
XX Claim 3; Page -; 83pp; English.
XX The invention relates to a novel method of detecting Mycobacterium
XX tuberculosis. The method comprises amplifying the DNA in the samples to
XX generate a detectable amount of amplified DNA comprising a catalase-
XX peroxide (katG) DNA fragment with sequence of bases 904-1523 of the
XX M. tuberculosis strain H37Rv MC katG gene sequence. The method optionally
XX further comprises determining if the katG DNA fragment has a serine to
XX threonine mutation in codon 315 (S315T mutation), indicative of an
XX isonicotinic acid hydrazide (INH)-resistant M. tuberculosis strain. The
XX method can be used to detect M. tuberculosis in biological fluids,
XX especially human sputum, useful to diagnose tuberculosis. This disease is
XX a major cause of human morbidity and mortality, and conclusive diagnosis
XX and subsequent treatment depends on identification of the etiologic agent
XX M. tuberculosis. INH has been used in tuberculosis treatment, but
XX INH-resistant strains have emerged; the method allows such drug-resistant
XX strains to be identified. The present sequence represents the M.
XX tuberculosis katG gene variant.
XX Note: this sequence is not provided in the specification; it has been
XX created by modifying the katG gene sequence given in Fig 7.
XX Sequence 2331 BP; 455 A; 735 C; 762 G; 379 T; 0 other;
Query Match 19.9%; Score 296.8; DB 20; Length 2331;
Best Local Similarity 62.3%; Pred. No. 1.9e-85;
Matches 501; Conservative 0; Mismatches 297; Indels 6; Gaps 2;
QY 688 ATCCCTGGGGGCTGATTTGATTATGATCCACAGATTTCACAGCTGGATGAGGCTC 747
DB 236 ACCCGATGGTGGCGGTTCGATATGCGCGGAGTCCGACCATCGACGTTGACGCC 295
QY 748 TGAATAAGATATCAAGATTGCTGACAACTCCAGGATTTGTCGCCCTCGCGGATATG 807

Db 296 TGACGGGACATCGAGGAAGTATGACCACTCGCAGCGTGGTGGCCGCGGACTACG 355
 QY 808 GTCAATATGATGCTCTTTCTTATTCGATGCGTGGCAGCGTGGCCGAACATACAGACAT 867
 Db 356 GCACATACGGGCGGCTTTATCCGATGCGTGGCAGCGTGGCCGACTACCGCATCC 415
 QY 868 ATGATGGCGGGAGCGGCGAGTGGTGCAGCAAGCTTTGAACCGCTGAACGCTGGC 927
 Db 416 ACACGCGCGCGCGCGCGCGCGGCGCATGACGCGGTTTCGCGCGCTTAACAGCTGGC 475
 QY 928 CGGATAACGTTATCTGGATAAAGCCCGCTGATTTGCTGGCCAGTCAAGAAAAATACG 987
 Db 476 CGGACAACGCCAGCTTGGACAAGCGCGCGCGCTGCTGGCGGCTCAAGAAGTACG 535
 QY 988 GCTCCAGTATTTCTGGGAGACCTGATGCTGCTGATGCTATGTTATGTTGCTGCTGAATCA 1047
 Db 536 GCAAGAAGCTCTCATGGCGGACCTGATTTGTTTCGCGGCACTGCGGCTGGATCGA 595
 QY 1048 TGGGATTTAAACGCTGGGATTTGCTGGCGGAAGAGATGACTGGAGTCGGACCTGG 1107
 Db 596 TGGGCTTCAGACGTTTCGGGTTTCGGCTTCGGCGGCTCGACAGTGGGAGCCGATGAGG 655
 QY 1108 TATACATGGGCGGCTGACAAACGCTTTTCAGATAACCGGGATA---AAAACGGGAAC 1164
 Db 656 TCTATTGGGCGAAGGACCACTGCTGCGGATGACGCTTACAGCGGTAAACCGGATC 715
 QY 1165 TTCAAGAACTCTTGGCGGACCGACAGATGGGACTTATTTATGTCATCTGTAAGGCCCG 1224
 Db 716 TGGAGAACCCTGCTGGCGGCGGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 775
 QY 1225 GTGGAACACAGATCTCTGCTTCGCGGAAGATATCAGGGAAGCTTTTTCAGTATGG 1284
 Db 776 ACGGAACCGGACCCCATGCGCGCGGCTGACATTCGCGAGAGCTTTTCGGCGCATGG 835
 QY 1285 CCATGATGATGAGGAGCTGTGCTGCTGATCGCGGAGGCGCATACATTTGTTAAAGCAC 1344
 Db 836 CCATGAACGACGTCGAACAGCGCGCTGATCTGCGCGTCACTTTTCGGTAAGACCC 895
 QY 1345 ATGGTGCAGGCTCTCTGAAATATGATTTGGCGCAGGCGCTGATGTCACCTGTGGAGG 1404
 Db 896 ATGGCGCG---GCCCGCGCGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 952
 QY 1405 AGCAGGAGCTGGGATGAAATATGATGTTGATGATGATGATGATGATGATGATGATGAT 1464
 Db 953 AGATGGCTTGGCTGGAAGAGCTGATGATGATGATGATGATGATGATGATGATGATGAT 1512
 QY 1465 GTGGCTTGAAGAGCTGCTGCTGA 1488
 Db 1013 CCGGCATCGAGGTCGTATGACGA 1036

RESULT 9
 AAT89968
 ID AAT89968 standard; DNA: 2221 BP.

XX AC AAT89968;

XX DT 20-MAR-1998 (first entry)

XX Mycobacterium tuberculosis strain H37Rv variant partial katG gene.

XX Tuberculosis; katG gene; isoniazid resistance; INH;

XX Isonicotinic acid hydrazide; ss.

XX Mycobacterium tuberculosis.

XX Key Location/Qualifiers
 CDS 1..2217

XX /tag= a

XX /product= "katG gene"

XX /note= "partial cds"

XX US5688639-A.

XX 18-NOV-1997.
 PD XX
 PF 18-APR-1994; 94US-0228662.
 XX XX
 PR 18-APR-1994; 94US-0228662.
 XX XX
 PA (MAYO-) MAYO FOUNDATION.
 XX XX
 PI Cockerill FR, Kline BC, Uhl JR;
 XX WPI; 1998-007975/01.
 XX
 PT Determination of isoniazid sensitivity of Mycobacterium tuberculosis
 PT strains - by restriction length polymorphism analysis of katG gene
 PS Disclosure; Column 12-14; 18pp; English.
 XX
 CC This partial DNA sequence encodes the katG gene from a variant
 CC strain of Mycobacterium tuberculosis H37Rv which is used in a novel
 CC method to rapidly identify strains of M. tuberculosis which are
 CC resistant to isoniazid (INH, also known as isonicotinic acid hydrazide).
 CC The method involves the use of restriction fragment length polymorphism
 CC (RFLP) analysis to determine if a NciI-MspI restriction site is absent
 CC in the DNA of the strain at the codon corresponding to codon 463 of a
 CC M. tuberculosis katG gene consensus sequence (see AAW31343). The absence
 CC of the site indicates an INH-resistant strain.
 XX
 SQ Sequence 2221 BP; 432 A; 691 C; 735 G; 363 T; 0 other;
 Query Match 18.4%; Score 274.2; DB 19; Length 2221;
 Best Local Similarity 61.4%; Pred. No. 4.3e-78;
 Matches 493; Conservative 0; Mismatches 303; Indels 7; Gaps 3;
 QY 688 ATCCCTGGGGGCTGATTTGATTATGCCACCAAGATTTCAACAGCTGGATATGAGGCTC 747
 Db 176 ACCCGATGGTGGGGCTTCGACTATGCCGCGAGGTCGCGGACCATGCTGACGCCC 235
 QY 748 TGAATAAAGATATCAAGATTTGCTGACAACTTCCAGGATTTGGTCCCTCGCGATATG 807
 Db 236 TGACGGGACATCGAGGAAGTATGATACCACTCGCAGCGTGTGGCCGCGGACTACG 295
 QY 808 GTCAATATGCTCTTTCTTATTTGATGCTTGGCAGCTGGCCGGAACATACAGGACAT 867
 Db 296 GCACATACGGGCGGCTTTTATCCGATGCGCTGGCAGCTGCGGCGCATCCGCGATCC 355
 QY 868 ATGATGCGGGGAGCGCGCAGTGGTGTGTCAGCAACGTTTGAACCGCTCAACAGCTGGC 927
 Db 356 ACACGCGCGCGCGCGCGCGCGGCGGCGCATGACGCGTTCGCGCGCTTAACAGCTGGC 415
 QY 928 CGGATAACGTTATCTGGATAAAGCCCGCTGATTTGCTGGCCAGTCAAGAAAAATACG 987
 Db 416 CGGACAACGCCAGCTTGGACAAGCGCGCGCGCTGCTGTCGCGGCTCAAGAAGTACG 475
 QY 988 GCTCCAGTATTTCTGGGAGACCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1045
 Db 476 GCAAGAAGCTCTCATGGCGGACCTGATTTGTTTCGCGGCAACCGCTGCGCTCGGAATC 535
 QY 1046 CATGGGATTTAAACGCTGGGATTTGCTGGCGGAAGAGATGACTGGGAGTCGGACCT 1105
 Db 536 GATGGCTTCAGACGCTTGGGTTTCGGCTTCGG--CGGTCAGCAGTGGGAGACCGATGA 593
 QY 1106 GGTATATGGGGCCTGACAAAGCTCTTTCAGATTAACCGGGATAAAAAACGGGAACCT 1165
 Db 594 GGTCTATTGGGCAAGGAAGCCACCTGGCTCGCGGATGACGGTTACAGCGTAAGCGATCT 653
 QY 1166 TCAGAAACCTCTTTCGCGCCACGAGATGGGACTTATTTATGCTCAATCCTGAAGCCCGG 1225
 Db 654 GGAGAACCCGCTGCGCGGCTGAGATGGGCTGATCTAGCTGAACCCGAGGCGCGGAA 713
 QY 1226 TGGAAAACCATGATCTCTGCTTCCGCGAAGATATCAGGGAAGCTTTTTCAGTATGCG 1285
 Db 714 CGGCAACCCGACCCCATGGCGGCGGTTCGACATTCGCGGAGACGTTTTCGGCGCATGGC 773

XX DT 11-MAY-1999 (first entry)
 XX DE M. tuberculosis katG gene.
 XX KW katG; catalase; peroxidase; enzyme; isonicotinic acid hydrazide;
 XX KW isoniazid; INH; diagnosis; resistance; ss.
 XX OS Mycobacterium tuberculosis.
 XX FH Key Location/Qualifiers
 XX FT CDS 1979..4186
 XX FT /*tag= a
 XX FT /gene= "katG"
 XX FT /product= "catalase-peroxidase"
 XX PN US5871912-A.
 XX PD 16-FEB-1999.
 XX PF 02-JUN-1995; 95US-0459499.
 XX PR 02-JUN-1995; 95US-0459499.
 XX PR 30-APR-1992; 92US-0875940.
 XX PR 14-AUG-1992; 92US-0929206.
 XX PR 11-MAR-1993; 93US-0029655.
 XX PA (INSP) INST PASTEUR.
 XX PI Cole ST, Heym B, Young DB, Zhang Y;
 XX DR WPI; 1999-166622/14.
 XX DR P-PSDB; ANW78363.
 XX PT New nucleic acid encoding catalase-peroxidase HPI - useful for rapid
 XX PT identification of isoniazid resistant Mycobacterium tuberculosis
 XX PS Disclosure; Fig 10A; 46pp; English.
 XX CC This sequence represents the Mycobacterium tuberculosis katG gene, which
 XX CC encodes a catalase-peroxidase enzyme isolated from a M. tuberculosis
 XX CC strain resistant to isonicotinic acid hydrazide (isoniazid or INH).
 XX CC The sequence is useful in the diagnosis of tuberculosis caused by
 XX CC INH-resistant M. tuberculosis and in the identification of other
 XX CC INH-resistant strains.
 XX SQ Sequence 4795 BP; 892 A; 1532 C; 1559 G; 812 T; 0 other;
 Query Match 18.4%; Score 274.2; DB 20; Length 4795;
 Best Local Similarity 61.4%; Pred. No. 6.9e-78;
 Matches 493; Conservative 0; Mismatches 303; Indels 7; Gaps 3;
 QY 688 ATCCCTGGGGGCTGATTGTTATATGCCACACAGATTTCACACAGCTGGATATGAGGCTC 747
 DB 2145 ACCCGATGGTGGCGGCTTCGACTATGCCGCGAGGTTCGGACCACTGCGACCTC 2204
 QY 748 TGAATAAGATATCAAGATTCTGCTGACAACTTCCACAGGATGTCGCCCTGCGGATATG 807
 DB 2205 TGACGGGGACATCGAGGAGTATGATGACCACTCGCAGCGTGTGTCGCGGCTACG 2264
 QY 808 GTCATATGTCCTTCTTTATTCGTATGCTTGGCAGCGTGGCGAACAATACAGGACAT 867
 DB 2265 GCCACTACGGCGCTGTTATCCGGATGGCGTGGCAGCTGCCGCACTCCCGCATCC 2324
 QY 868 ATGATGCCGGGGAGCGCGAGTGGTGTGTCAGCAAGCTTTTGAACGCTGAACAGCTGGC 927
 DB 2325 ACCAGCGCGCGCGCGCGCGCGGCGGATGATGACGGTTCGCGCGCTTAACAGCTGGC 2384
 QY 928 CGGATACGTTATCTGATTAAGCCCGCTGATGTCGTGGCAGTCAAGAAATAATACG 987
 DB 2385 CCGACACGCGAGCTTGGACAAGCGCGCGGCTGTGTGCGCGGTCAAGAAGATACG 2444
 QY 988 GCTCCAGTATTTCTGGGGAGACCTGATGCTCTGACTGTAA--TGTTGCCCTTGATC 1045

DB 2445 GCAGAAGCTCTCATGGGGGACCTGATGTTTTCGCCGCAACCGCTGCGCTCGGAATC 2504
 QY 1046 CATGGGATTAAACGCTGGGATTTCGCTGCGGAAGAGAGATGACTGGGAGTGGACCT 1105
 DB 2505 GATGGGCTTCAAGACGCTTCGGGTTTCGGCTTCGG--GCGTCGACCACTGGGAGACCGATGA 2562
 QY 1106 GGTATACTGGGGGCTTGACAAAGCTCTTTCGAGATAACCGGGATAAAGCGGAACCT 1165
 DB 2563 GGTCTATTGGGGCAAGGAAGCCACCTGGCTTCGGCGATGACGGTTACAGCGTAAGCGATCT 2622
 QY 1166 TCAGAAACCTCTTTCGCCGCCACGACAGATGGGACTTATTATGTCAATCTCAAGGCCCGG 1225
 DB 2623 GGAGAACCCGCTGCGCGGTGCAGATGGGCTTGATCTAGTGAACCCGGAGCGCGGAA 2682
 QY 1226 TGGAAACCCAGATCTCTGCTGCTCCGCGAAAGATATCAGGAAGAGCTTTTTCAGCTATGC 1285
 DB 2683 CGGCAACCCGACCCCATGCGCGGCTCGACATTCGGGAGACGTTTCGGCGCATGGC 2742
 QY 1286 CATGGATGATGAGGAGACTGTGCCCTGATCGCGGGAGGCGCATATTTGGTAAAGACA 1345
 DB 2743 CATGAACGAGCTCGAAACAGCGGCTGATCGTCGCGGCTCACACTTTTCGGTAAGACCA 2802
 QY 1346 TGGTGCAGCGCTCTCTGAAATGATTTGGCGCAGGCGCTGATGTCACCTCTGGAGGA 1405
 DB 2803 TGGCGCGG---GCCCGCGGATCTGTCGCGCCCGCAACCGGAGGCTGCTCCGCTGGAGCA 2859
 QY 1406 GCAGGAGCTGGGATGGAATAATTAATGTCGTACAGGAACCGCAATATACCATCACCAG 1465
 DB 2860 GATGGCTTGGGCTGGAAGAGCTCGTATGSCACCGCAACCGGTAAAGACCGCATCACCAG 2919
 QY 1466 TGGCCTGGAGGAGCGCTGGTCGA 1488
 DB 2920 CGGCATCGAGGTCGTATGGACGA 2942
 RESULT 12
 AAQ06815
 ID AAQ06815-standard; DNA; 2196 BP.
 AC AAQ06815;
 DT 06-MAR-1991 (first entry)
 XX Sequence encoding heat resistant peroxidase.
 DE pod10; E.coli UM228; ds.
 XX Bacillus stearothermophilus.
 XX Key Location/Qualifiers
 FT CDS 1..2193
 FT /*tag= a
 XX JP02268684-A.
 XX 02-NOV-1990.
 XX 07-APR-1989; 89JP-0089469.
 XX 07-APR-1989; 89JP-0089469.
 XX (TOYM) TOYOBO KK.
 XX WPI; 1990-372008/50.
 XX P-PSDB; AAR08205.
 XX DNA contg. genetic information of heat resistant peroxidase -
 XX prepd. by culturing transformant and collecting heat resistant
 XX peroxidase
 XX Disclosure; Fig 2; 8pp; Japanese.
 XX

Db 458 CTGGCCGACACCTCAGCTGGATTAAGCCGCGCTGCTGCTGGCCGATCAAGA 517
 QY 982 AATACGGCTCCAGTATTCTCTGGGAGACCTGATGCTCTGACTGTAATGTTGCCCTTG 1041
 Db 518 AGTACGGCAACAAATCAGCTGGCAGACCTGATGATTTCTGGCTGGCAGCGTGGCTATG 577
 QY 1042 AATCCATGGGATTTAAACGCTGGGATTTCTGGCGGAAGAGATGACTGGGACTCGG 1101
 Db 578 AGTCCATGGCTTACCTGCTTACGGCTTCTTTCTGGCGCGCTGCGATATTTGGGAACCCG 637
 QY 1102 AACTGG---TATCTGGGGCTTGACAAACAGCTCTTTCAGAT----- 1142
 Db 638 AAAAGATATCTACTGGGTGACGAAAAGAGTGGCTGGCACCTTCTGACGACGTACG 697
 QY 1143 --AACCGGATAAAGCGGAACCTTCAGAAACCTCTTGGCGCCGACGAGATGGGACTTA 1200
 Db 698 GCGACGTGAACAAAGCCAGAGACCATGGAACCCGCTGGCGGCTGTCCAAATGGGTCTGA 757
 QY 1201 TTTATGTCATCTGAAGCCCGCTGGAAACCCAGATCTCTGGCTTCCGCGAAAGATA 1260
 Db 758 TCTATGTGACCCGGAGGTGTTAAGCGCCACCTGATCGCTGAGAACCGCACGAGG 817
 QY 1261 TCAGGAAGCTTTTTCACGTATGGCCATGATGAGGAGACTGTGGCCCTGATCGCG 1320
 Db 818 TACTTGAACCTTTCGCGCTATGGCGATGAACGACGAAACCCGAGCCCTCACAGCTG 877
 QY 1321 GAGGCGATCATTTGTTAAGCACATGTTGACGCTCTCTGAAATGTTATTTGGCGAG 1380
 Db 878 GCGGCCACACCGTCGCTGATGTCACGGTAATGGCAATGCTCTGCTGCTGCTG 934
 QY 1381 GGCTGATGTGACCTGTGGAGGACGAGGACTGGGATGGAAATAAATGTTGGTACAG 1440
 Db 935 ACCAAAGCTCTGACGTGTAACACGAGGCTTAGTGGGCAACCCCAACATGACGAG 994
 QY 1441 GAAACGGCAATATACCATCACAGTGGCTTGAAGGAGCGCTGGTGGAC 1489
 Db 995 GCAAGCAAGCAACGCGTGACCTCGGGTATCGAAGGTGCTTGGACAC 1043

RESULT 14

AA49864
 ID AAA49864 standard; DNA; 660 BP.
 AC AAA49864;
 XX
 DT 25-SEP-2000 (first entry)
 XX
 DE Mycobacterium tuberculosis katG gene (isoniazid resistance).
 XX
 KW Antibiotic resistance; katG gene; isoniazid resistance; ss.
 XX
 OS Mycobacterium tuberculosis.
 XX
 FH Key Location/Qualifiers
 FT primer_bind complement(41..60)
 FT /*tag- a
 FT /note- "primer of AAA49827"
 FT 231..250
 FT /*tag- b
 FT /note- "primer of AAA49828"
 XX
 PN WO200036142-A1.
 XX
 PD 22-JUN-2000.
 XX
 PF 10-DEC-1999; 99WO-CA01177.
 XX
 PR 11-DEC-1998; 98US-0111794.
 XX
 PA (VISI-) VISIBLE GENETICS INC.
 XX
 PI Shipman R;
 XX

DR XX WPI; 2000-431611/37.
 PT Method for the detection and characterization of Mycobacterium
 XX tuberculosis with antibiotic resistance in a sample -
 PS Disclosure; Page 5-6; 43pp; English.
 CC The present sequence is that of the Mycobacterium tuberculosis
 CC katG (isoniazid resistance) gene (bp661-1320). Amplification and
 CC cycle sequencing primers (see AAA49823-62) are used for the detection
 CC and analysis of antibiotic resistance-associated mutations in
 CC defined regions of rpoB (rifampin), katG (isoniazid), oxyR-aphC PR
 CC (isoniazid), mbaA (isoniazid), rpsL/s12 (streptomycin), 16S/rfs
 CC (streptomycin), emdB (ethambutol), pncA (pyrazinamide), gyrA
 CC (ciprofloxacin) and 23S (azithromycin) genes of M. tuberculosis.
 CC These primers can be used in a method for the detection and
 CC characterization of M. tuberculosis present in a sputum sample.
 CC The method involves performing a sequencing procedure, with or
 CC without prior amplification, to detect the presence of M.
 CC tuberculosis, and if present to evaluate the rpoB, katG, rpsL/s12
 CC and 23S genes for the presence of antibiotic-inducing mutations.
 CC If M. tuberculosis is detected, a second sequencing procedure is
 CC performed on the sample to evaluate additional genes for the
 CC presence of antibiotic resistance-inducing mutations. Genotypic
 CC tests are rapid, sensitive and accurate providing information as to
 CC antibiotic treatment options.
 XX
 SQ Sequence 660 BP; 127 A; 206 C; 227 G; 100 T; 0 other;
 Query Match 8.2%; Score 122.8; DB 21; Length 660;
 Best Local Similarity 62.6%; Pred. No. 4.9e-29;
 Matches 209; Conservative 0; Mismatches 122; Indels 3; Gaps 1;
 QY 1155 AAGCGGAACCTTCAGAAACCTCTTGGCCGACGACGATGGGACTTATTTATGTCATCCT 1214
 Db 26 AAGCGGGATCTGGAGAACCGCTGCGCGGTCAGATGGGCTGATCTACGTGAACCG 85
 QY 1215 GAAGCGCCGCTGGAAACACGATCTCTGGCTTCCGCGAAGATATCAGGGAAGCTTTT 1274
 Db 86 GAGGGGCGAACGACGCAACCGGACCCATGCGCGCGGCTGACATTCGCGAGAGCTTT 145
 QY 1275 TCACGTATGGCCATGATGATGAGGACGCTGTCGCTGATCGCGGAGGACATACATTT 1334
 Db 146 CGCGCGATGCGCCATGACGACGTCGAAACACGCGCGCTGATCTGCGCGCTCACATTC 205
 QY 1335 GGTAAAGCACATGGTGCAGCGCTCTCTGAAATAATGTTGGCGCAGCGCTGATGGTGCA 1394
 Db 206 GGTAAAGCACATGGCGCGG---GCCGCGCGATCTGTCGCGCCCGAACCCGAGGCTGCT 262
 QY 1395 CTTGTGGAGGACGAGGACTGGGATGGAAATAATGTTGGTACAGGAACCGCAATAT 1454
 Db 263 CCGCTGGAGCAGATGGGCTTGGCTGGAAGAGCTCGTATGGCACCGACCGGTAAGGAC 322
 QY 1455 ACCATCACCAGTGGCCTGGGAAGGAGCTGCTGTCGA 1488
 Db 323 GCGATCACCAGCGGCATCGAGTCTGATGGACGA 356
 RESULT 15
 AAF07490
 ID AAF07490 standard; cDNA; 1852 BP.
 AC AAF07490;
 XX
 DT 13-MAR-2001 (first entry)
 XX
 DE Fusarium venenatum EST SEQ ID NO:13.
 XX
 KW Multiple gene expression; filamentous fungal cell; EST;
 KW expressed sequence tag; Fusarium venenatum; Aspergillus niger;
 KW Aspergillus oryzae; Trichoderma reesei; Identification; recombination;
 KW culture condition; environmental stress; spore morphogenesis;
 KW metabolic pathway engineering; catabolic pathway engineering; ss.

Db 25 CCGGTACGCGGCTAGAAAATTCGGTTTCGTAATTCGGTTTGGTGTGGTGGTG 84
 Qy 1084 AAGATGACTGGGAGTCGGACCTGGTATACCTGAGGAGGCTTTCGAGATA 1143
 Db 85 AAGACGTTTGGGAATCGGATTTGGATGTTAATTTGGGCTGATGAAAAGTTTGGTGTATT 144
 Qy 1144 ACCGGGATAAAGCGGAATTCAGNAACCTCTTCCGCCACGAGATGGGACTATTT 1203
 Db 145 ATCGTTATTCGGAAGCGTTGGCGAAAGTATCGTTGGGTGTAATTCGAGATGGGTTGATT 204
 Qy 1204 ATGTCAATCTGAAGCCCGGTTGGAAACAGATCTCTGCTTCGCCGAAAGATATCA 1263
 Db 205 ACCTTAATTCGGAAGTTCCGATATAGCGCGAATCGTTTTCGGTAGTAGTATTC 264
 Qy 1264 GGAAGCTTTTACGATGATGCGGATGATGAGGAGACTGTGGCCCTGATCGCGGAG 1323
 Db 265 GCGGATTTTCGGTAATATGGTATGATGACGACGAAAGAAATCGTGGCGTTGATTGCGGGTG 324
 Qy 1324 GGCATACATTTGTAAGCACATGTCACGCTCTCTGAAATATGTTGGCGCAGGC 1383
 Db 325 GTTATAGTTGGGTAAATTTACGGTGTCTGGTTCGATATTTAA--TGTAGGTTTGTATT 381
 Qy 1384 CTGATGTCACCTGTGGAGGACGAGGACTGGGATGGAATAAATGTTGGTACAGGAA 1443
 Db 382 TAGAGTTGATCGATTGAAGATAGGTTTAGTTGGGCGAGTATTTACGGTAGCGGCG 441
 Qy 1444 ACGGCAATATACATACACAGTGGCTGGAAGGAGCCTGG 1484
 Db 442 TTGCGTAGATGTTATTTATTTTGGTTTGGTGAAGTAGTTGG 482

RESULT 17

ABQ23235/c
 ID ABQ23235 standard; DNA; 538 BP.

XX AC ABQ23235;
 XX DT 12-JUL-2002 (first entry)
 XX OLigonucleotide for detecting cytosine methylation SEQ ID NO 9826.
 DE Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;
 KW drug; side effect; cancer; central nervous system; cardiovascular;
 KW gastrointestinal; respiratory system; single nucleotide polymorphism;
 KW SNP; cell differentiation; ds.
 XX OS Homo sapiens.
 XX WO200218632-A2.
 XX PD 07-MAR-2002.
 XX 01-SEP-2001; 2001WO-EPI0074.
 XX 01-SEP-2000; 2000DE-1043826.
 XX 05-SEP-2000; 2000DE-1044543.
 XX (EPIC-) EPIGENOMICS AG.
 XX Olek A, Piepenbrock C, Berlin K, Guetig D;
 XX WPI; 2002-371829/40.
 XX Determining the degree of cytosine methylation in genomic DNA, useful
 PT for diagnosis and prognosis, comprises selective hybridization of
 PT amplicons from chemically treated DNA -
 XX Claim 12; 56pp + Sequence Listing; 56pp; German.
 XX This invention describes a novel method for determining the degree of
 CC methylation of a particular cytosine in a motif 5'-CpG-3', present in a
 CC genomic sample of DNA. The sample is treated chemically to convert

CC cytosine (C) but not methylated C, to uracil, then part of the genomic
 CC DNA that contains the target C is amplified to form a labeled amplicon.
 CC The amplicon is hybridised to two classes, each with at least one
 CC member, of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers
 CC and the degree of hybridisation to both classes is determined from the
 CC label on the amplicon. From the ratio of labels hybridised to the two
 CC classes of oligomers, the degree of methylation is calculated. The method
 CC is used: (i) for diagnosis and/or prognosis of side effects of
 CC therapeutic drugs and of a wide range of diseases, e.g. cancer, disorders
 CC of the central nervous, cardiovascular, gastrointestinal and respiratory
 CC systems etc., particularly by detecting mutations or single nucleotide
 CC polymorphisms (SNP's); and (ii) for differentiation of cell or tissue
 CC types and for investigating cell differentiation. The method allows the
 CC methylation status of many C residues to be determined simultaneously.
 CC ABQ13410-ABQ54121 represent genomic DNA sequences used to illustrate the
 CC method for determining the degree of cytosine methylation described in
 CC the disclosure of the invention.

SQ Sequence 538 BP; 206 A; 166 C; 47 G; 119 T; 0 other;

Query Match 6.18; Score 91.4; DB 24; Length 538;
 Best Local Similarity 51.48; Pred. No. 7.5e-19;
 Matches 237; Conservative 0; Mismatches 221; Indels 3; Gaps 1;
 Qy 1024 CTGTAATGTTCCTTGAATCCATGGATTTAAACGCTGGGATTTGCTGGCGGAAGAG 1083
 Db 514 CCGGTAACTGGGCGTTAGAAAATTCGGTTTTCGTAATTTTCGGTTTGGTGTGGTGGT 455
 Qy 1084 AAGATGACTGGGAGTCGGACCTGTATACCTACCTGAGGAGGCTGACAAACGCTCTTCACAGATA 1143
 Db 454 AAGACGTTTGGGATCGGATTTGGATGTTAATTTGGGCTGATGAAAAGTTTGGTGTATT 395
 Qy 1144 ACCGGGATAAAGCGGAATTCAGAAACCTCTTCCGCCACGAGATGGGACTTATTT 1203
 Db 394 ATCGTTATTCGGAAGCGTTGGCGAAAGTATCGTTGGGTGTAATTCGAGATGGGTTGATT 335
 Qy 1204 ATGTCAATCTGAGGCGCCGCTGGAACACAGATCCTCTGGCTTCCGCGAAGATATCA 1263
 Db 334 ACGTAAATTCGGAAGGTTCCGATTAAGCGCGGAATCGTTTTCGGGTAGTAGTTATTC 275
 Qy 1264 GGAAGCTTTTTCACGTATGCGCATGATGATGAGGAGACTGTGGCCTCATCGCGGAG 1323
 Db 274 GCGCATTTTCGGTAAATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 215
 Qy 1324 GGCATACATTTGGTAAAGCACATGCTGACGCTCTCTGAAATATGTTGGCCGAGGC 1383
 Db 214 GTTATACGTTGGGTAAATTTTACGCTGCTCGGTTTCGATATTTAA--TGTAGGTTTGTATT 158
 Qy 1384 CTGATGCTGACCTGTGAGGAGCAGGAGTGGGATGGAATAAATATGTTGGTACAGAA 1443
 Db 157 TAGAAGTTGATCGATTGAAGATAGGTTTAGGTTGGGCGAGTATTTACGGTAGCGGCG 98
 Qy 1444 ACGGCAATATACCATCACCAGTGGCTGGAAGGAGCCTGG 1484
 Db 97 TTGCGGTAGATGTTATTTATTTTGGTTTGGGAGTAGTTTGS 57

RESULT 18

AAF95160
 ID AAF95160 standard; DNA; 150 BP.

XX AC AAF95160;
 XX DT 23-MAY-2001 (first entry)
 XX DE katG gene probe #2.

XX Tubercle bacillus; drug sensitivity; drug resistance; rifampicin;
 KW streptomycin; kanamycin; isoniazid; ethambutol; rpoB gene; rrs gene;
 KW rpsL gene; inhA gene; katG gene; embB gene; probe; PCR primer; ss.
 XX OS Mycobacterium tuberculosis.

```

PN EP1076099-A2.
PD 14-FEB-2001.
PF 02-AUG-2000; 2000EP-0306563.
PR 03-AUG-1999; 99JP-0220357.
XX (NISN ) NISSHINBO IND INC.
PA (SYST-) SYSTEM RES INC.
XX
PI Suzuki Y, Nishida M, Takenishi S;
XX WPI; 2001-246696/26.
XX
XX New oligonucleotides, nucleic acid probes and primers are useful for
PT differentiating drug-resistance and determining infection with tubercle
PT bacilli -
XX
PS Disclosure; Page 61-62; 114pp; English.
XX
XX The present invention relates to oligonucleotides based on nucleotide
CC sequences obtained from both wild-type tubercle bacilli (wtTB) that are
CC susceptible to a drug and mutant-type tubercle bacilli (mtTB) that are
CC resistant to a drug. The drugs used in the present invention are
CC rifampicin (RFP), streptomycin (SM), kanamycin (KM), isoniazid (INH) and
CC ethambutol (EB). The rpoB gene is responsible for resistance to RFP; the
CC rrs gene is responsible for resistance to SM and KM; the rpsL gene is
CC responsible for resistance to SM; the inhA gene is responsible for
CC resistance to INH; the katG gene is responsible for resistance to INH;
CC and the embB gene is responsible for resistance to EB. The present
CC invention also relates to nucleic acid probes having part of a nucleotide
CC sequence of tubercle bacilli (TB) responsible for drug resistance and
CC primers used to generate the probes. The present sequence is an
CC oligonucleotide of the present invention. The oligonucleotides of the
CC present invention can be used to enable the differentiation of drug
CC resistance and the determination of infection with tubercle bacilli
CC simultaneously.
XX
SQ Sequence 150 BP; 29 A; 51 C; 53 G; 17 T; 0 other;

Query Match 5.1%; Score 75.6; DB 22; Length 150;
Best Local Similarity 71.7%; Pred. NO. 4.9e-14;
Matches 99; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY 852 GGAACATACAGACATATGATGCGCGGGAGCGCCAGTGTGTCTACGACGTTTGA 911
DB 1 GGCACCTACCGATCCAGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 60
QY 912 CGCTGAACAGCTGCCGGATAACGTTAATCTGGATAAGCCGCTGATTGCTGCGCA 971
DB 61 CGGCTTACAGCTGGCGCGACACGACGCTTGGACAAGCGCGCGCGCGCGCGCG 120
QY 972 GTCAGAAAAAATACGGC 989
DB 121 GTCAGAGAAGTACGGC 138

RESULT 19
AAF95161
ID AAF95161 standard; DNA; 150 BP.
XX
AC AAF95161;
XX
DT 23-MAY-2001 (first entry)
XX
DE katG gene probe #3.
XX
KW Tubercle bacillus; drug sensitivity; drug resistance; rifampicin;
KW streptomycin; kanamycin; isoniazid; ethambutol; rpoB gene; rrs gene;
KW rpsL gene; inhA gene; katG gene; embB gene; probe; PCR primer; ss.
XX Mycobacterium tuberculosis.
XX

```

```

XX EP1076099-A2.
XX 14-FEB-2001.
XX 02-AUG-2000; 2000EP-0306563.
XX 03-AUG-1999; 99JP-0220357.
XX (NISN ) NISSHINBO IND INC.
XX (SYST-) SYSTEM RES INC.
XX
XX Suzuki Y, Nishida M, Takenishi S;
XX WPI; 2001-246696/26.
XX
XX New oligonucleotides, nucleic acid probes and primers are useful for
PT differentiating drug-resistance and determining infection with tubercle
PT bacilli -
XX
PS Disclosure; Page 62; 114pp; English.
XX
XX The present invention relates to oligonucleotides based on nucleotide
CC sequences obtained from both wild-type tubercle bacilli (wtTB) that are
CC susceptible to a drug and mutant-type tubercle bacilli (mtTB) that are
CC resistant to a drug. The drugs used in the present invention are
CC rifampicin (RFP), streptomycin (SM), kanamycin (KM), isoniazid (INH) and
CC ethambutol (EB). The rpoB gene is responsible for resistance to RFP; the
CC rrs gene is responsible for resistance to SM and KM; the rpsL gene is
CC responsible for resistance to SM; the inhA gene is responsible for
CC resistance to INH; the katG gene is responsible for resistance to INH;
CC and the embB gene is responsible for resistance to EB. The present
CC invention also relates to nucleic acid probes having part of a nucleotide
CC sequence of tubercle bacilli (TB) responsible for drug resistance and
CC primers used to generate the probes. The present sequence is an
CC oligonucleotide of the present invention. The oligonucleotides of the
CC present invention can be used to enable the differentiation of drug
CC resistance and the determination of infection with tubercle bacilli
CC simultaneously.
XX
SQ Sequence 150 BP; 28 A; 37 C; 54 G; 31 T; 0 other;

Query Match 4.5%; Score 66.8; DB 22; Length 150;
Best Local Similarity 65.3%; Pred. NO. 3.6e-11;
Matches 98; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

QY 962 GCTGTGCCAGTCAAGAAAAATACGGCTCCAGTATTTCTCTGGGAGACCTGATGCTCT 1021
DB 1 GCTGTGCCGCTCAAGAGAAGTACGCGCAAGAGCTCTCATGGCGGACCTGATTTT 60
QY 1022 GACTGGTAATGTGCCCTTGAATCCATGGGATTTAAACGCTGGGATTTCTGGCGGGAAG 1081
DB 61 CGCGGCAACTGCGCGCTGGAATCGATGGGCTTCAAGACGTTGCGGTTCTCGCGG 120
QY 1082 AGAAGATGACTGGGAGTCGACCTGCTATA 1111
DB 121 GGTCGACAGTGGGAGCCCGCATGAGGTCTA 150

RESULT 20
AAF95159
ID AAF95159 standard; DNA; 150 BP.
XX
AC AAF95159;
XX
DT 23-MAY-2001 (first entry)
XX
DE katG gene probe #1.
XX
KW Tubercle bacillus; drug sensitivity; drug resistance; rifampicin;
KW streptomycin; kanamycin; isoniazid; ethambutol; rpoB gene; rrs gene;
KW rpsL gene; inhA gene; katG gene; embB gene; probe; PCR primer; ss.
XX

```

OS Mycobacterium tuberculosis.
 PN EP1076099-A2.
 XX
 XX 14-FEB-2001.
 XX
 XX 02-AUG-2000; 2000EP-0306563.
 XX
 XX 03-AUG-1999; 99JP-0220357.
 XX
 XX (NISN) NISHINBO IND INC.
 PA (SYST-) SYSTEM RES INC.
 PA
 XX Suzuki Y, Nishida M, Takenishi S;
 PI
 XX WPI; 2001-246696/26.
 DR
 XX
 XX New oligonucleotides, nucleic acid probes and primers are useful for
 PT differentiating drug-resistance and determining infection with tubercle
 PT bacilli -
 XX
 PS Disclosure; Page 61; 114pp; English.
 XX
 XX The present invention relates to oligonucleotides based on nucleotide
 CC sequences obtained from both wild-type tubercle bacilli (H37Rv) that are
 CC susceptible to a drug and mutant-type tubercle bacilli (H37Rv) that are
 CC resistant to a drug. The drugs used in the present invention are
 CC rifampicin (RFP), streptomycin (SM), kanamycin (KM), isoniazid (INH) and
 CC ethambutol (EB). The rpoB gene is responsible for resistance to RFP; the
 CC rrs gene is responsible for resistance to SM and KM; the rpsL gene is
 CC responsible for resistance to SM; the inhA gene is responsible for
 CC resistance to INH; the katG gene is responsible for resistance to INH;
 CC and the embB gene is responsible for resistance to EB. The present
 CC invention also relates to nucleic acid probes having part of a nucleotide
 CC sequence of tubercle bacilli (TB) responsible for drug resistance and
 CC primers used to generate the probes. The present sequence is an
 CC oligonucleotide of the present invention. The oligonucleotides of the
 CC present invention can be used to enable the differentiation of drug
 CC resistance and the determination of infection with tubercle bacilli
 CC simultaneously.
 XX
 SQ Sequence 150 BP; 26 A; 53 C; 48 G; 23 T; 0 other;
 Query Match 4.3%; Score 63.8; DB 22; Length 150;
 Best Local Similarity 66.2%; Pred. No. 3.4e-10;
 Matches 92; Conservative 0; Mismatches 47; Indels 0; Gaps 0;
 QY 733 TGGATATGGAGGCTCTGAAAAAGATATCAAGATTTGCTGCAACTTCCAGGATTGGT 792
 DB 12 TCGACGTTGACGCCCTGACCGGGACATCGAGGAAGTATGACACCTCGCAGCGTGGT 71
 QY 793 GCCCTCGGATTATGTCATATGTCCTTTCTTTATTCGTATGGCTTGGCAGCGTCCG 852
 DB 72 GCGCCGCGGACTACGGCCACTACGGCGCGCTGTTTATCCGGATGGCGTGCAGCGTCCG 131
 QY 853 GAACATACAGGACATATGA 871
 DB 132 GCACCTACCGCATCCACGA 150
 RESULT 21
 ABQ23236/c
 ID ABQ23236 standard; DNA; 538 BP.
 XX
 AC ABQ23236;
 XX
 DT 12-JUL-2002 (first entry)
 XX
 DE Oligonucleotide for detecting cytosine methylation SEQ ID NO 9827.
 XX Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;
 KW drug; side effect; cancer; central nervous system; cardiovascular;
 KW gastrointestinal; respiratory system; single nucleotide polymorphism;
 KW

KW SNP; cell differentiation; ds.
 OS Homo sapiens.
 XX
 PN WO200218632-A2.
 XX
 PD 07-MAR-2002.
 XX
 XX 01-SEP-2001; 2001WO-EP10074.
 PF
 XX 01-SEP-2000; 2000DE-1043826.
 PR
 PR 05-SEP-2000; 2000DE-1044543.
 XX
 XX (EPIC-) EPIGENOMICS AG.
 XX
 XX Olek A, Piepenbrock C, Berlin K, Gueutig D;
 PI
 XX WPI; 2002-371829/40.
 DR
 XX
 XX Determining the degree of cytosine methylation in genomic DNA, useful
 PT for diagnosis and prognosis, comprises selective hybridization of
 PT amplicons from chemically treated DNA -
 XX
 PS Claim 12; 56pp + Sequence Listing; 56pp; German.
 XX
 XX This invention describes a novel method for determining the degree of
 CC methylation of a particular cytosine in a motif 5'-CpG-3', present in a
 CC genomic sample of DNA. The sample is treated chemically to convert
 CC cytosine (C) but not methylated C, to uracil, then part of the genomic
 CC DNA that contains the target C is amplified to form a labeled amplicon.
 CC The amplicon is hybridized to two classes, each with at least one
 CC member, of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers
 CC and the degree of hybridization to both classes is determined from the
 CC label on the amplicon. From the ratio of labels hybridized to the two
 CC classes of oligomers, the degree of methylation is calculated. The method
 CC is used: (i) for diagnosis and/or prognosis of side effects of
 CC of the central nervous, cardiovascular, gastrointestinal and respiratory
 CC systems etc., particularly by detecting mutations or single nucleotide
 CC polymorphisms (SNP's); and (ii) for differentiation of cell or tissue
 CC types and for investigating cell differentiation. The method allows the
 CC methylation status of many C residues to be determined simultaneously.
 CC ABQ13410-ABQ54121 represent genomic DNA sequences used to illustrate the
 CC method for determining the degree of cytosine methylation described in
 CC the disclosure of the invention.
 XX
 SQ Sequence 538 BP; 111 A; 47 C; 142 G; 238 T; 0 other;
 Query Match 3.2%; Score 47.4; DB 24; Length 538;
 Best Local Similarity 45.9%; Pred. No. 0.00017;
 Matches 162; Conservative 0; Mismatches 191; Indels 0; Gaps 0;
 QY 1014 ATGCTCTGACTGTGTAATGTGCGCTTGAATCCATGGATTAAACGCTGGGATTGGCT 1073
 DB 524 ATTATCTCCGATACGCTAACGCTAAACAACTCCGACTTCGTTCTGATTTAAT 465
 QY 1074 GCGGGAAGAAGATGACTGGAGTGGACCTGGTATGCTGCGGGGCTGACAAACAGCCT 1133
 DB 464 ACCGATCGTAAAAACGCTCTAAAAACCGAATCTAAATATTAACTAAATAATAAAAAACC 405
 QY 1134 CTTGCAGATAACCGGGATAAAAACGGGAAACCTTCAGAAACCTTCGCGCCACCCAGATG 1193
 DB 404 TAACTAACTACCGCTCATCCGNAACGCTAACGAAACACCGCTTAATATCAACGGAATA 345
 QY 1194 GGACTTATTATGTCATCTGAAGCCCGGGTGGAAACACAGATCTCTGGCTTCGCGC 1253
 DB 344 AATCTAATTTACGTTAACCCGAAAAACCCGAATCAACACGACGACCGCTTCTACGACA 285
 QY 1254 AAGATATACGGAAGCTTTTTCACGTATGCCATGGATGATGAGGATGTGGCCCTG 1313
 DB 284 ACAATATCCGCGGACCTTTCGACACATATAACATAAACGACGAAAAACCGTAACGCTA 225
 QY 1314 ATCGCGGAGGCGCATACATTTGGTAAAGCATGGTGCACCGCTCTCTGAAAA 1366

XX Producing genetic parthenocarp in plants - with sequence encoding
PT auxin modulator controlled by promoter specific for the ovary
XX
XX Claim 7; Fig 9; 41pp; English.
XX
XX This sequence comprises a genomic clone of the tomato fruit
CC expressed proline-rich protein gene TPRP-F1, including the promoter
CC region. A claimed method for the production of genetic
CC parthenocarp in plants comprises: providing a cassette including
CC DNA coding for modulation of auxin effects in plants and a
CC promoter specific for the ovary between anthesis and early fruit
CC development; and introducing the cassette into a plant. Suitable
CC promoters include the TPRP-F1 promoter and an alternative TPRP-F1
CC promoter (see A8789279) that confers ovary and developing embryo
CC specificity. The method allows parthenocarp to be induced in
CC especially tomato plants without the need to apply (synthetic)
CC auxin or auxin inhibitor to each truss.
XX
XX Sequence 4320 BP; 1450 A; 844 C; 585 G; 1441 T; 0 other;

Query Match 2.6%; Score 38.8; DB 18; Length 4320;
Best Local Similarity 43.8%; Pred. No. 0.4;
Matches 169; Conservative 0; Mismatches 217; Indels 0; Gaps 0;

QY 424 CTTAGTATTTTGTCTGTAGTATCTATCCAGCAATAGGTATATCTGTTCATCAATA 483
DB 3628 CTTATGTATGATTTGGCAAGTATCCACCAAGACITCAAGTGTCTTCAACCTAAATC 3687

QY 484 AAGTTGACATTTGTATACAAATGCGAATTTCCCTTAATCCGGAGCTATTCGTATGATAA 543
DB 3688 AAGGTTTCCACTTTTTCACACTTTCAATTTATCTACTCTACTCAATTTATGTGTAC 3747

QY 544 AAAAATCTTCTCTGTTCTGATTTCTCTGCGCTATCGGAGCTTTTCTACCGCTGTAG 603
DB 3748 AGTTGACATTTCAAGTATAGCCCAATTTCTAGTCGGAATTTTAAATCTCTTT 3807

QY 604 CGCTGATAAAGAGACTCAAAATTTCTACTATCCAGAACACTGGATTTAACTCCTC 663
DB 3808 AAATATTTTGTATATCTACTTATAGTATTTTTCATAGTATAAATTTCAATTCATA 3867

QY 664 TCAGATTACACAGCCCTGAATCAATCCCTGGGGGCTGATTTGATTTATCCACCAT 723
DB 3868 TATGAATTCACGGTCAAAATTTTAAAGTTTATTTAAACCAATGCCACATAACTTAGAACAT 3927

QY 724 TTCACAGCTGGATATGGAGGCTCTGAAAAAGATATCAAGATTTGCTGACAACTTCCC 783
DB 3928 ACAATATCTTTTGTATCAAGATTTGGCAATTCGTATACATAATCTTTAGCAAGTATA 3987

QY 784 AGGATGGTGCCCTCGGATTAAGT 809
DB 3988 TGTATACCAACATTATGTAATATGAT 4013

RESULT 26
ABL08713
ID ABL08713 standard; cDNA; 5591 BP.
XX
XX ABL08713;
XX
XX 26-MAR-2002 (first entry)
XX
XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 20621.
DE Drosophila; developmental biology; cell signalling; insecticide;
XX
XX pharmaceutical; gene; ss.
XX
XX Drosophila melanogaster.
OS
XX WO200171042-A2.
XX
XX 27-SEP-2001.
XX

PF 23-MAR-2001; 2001WO-US09231.
XX
XX 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX
XX (PEKE) PE CORP NY.
XX
XX Venter JC, Adams M, Li PWD, Myers EW;
XX
XX WPI; 2001-656860/75.
DR P-PSDB; ABB64610.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
XX Claim 1; SEQ ID NO 20621; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

QY Sequence 5591 BP; 1876 A; 1209 C; 1357 G; 1149 T; 0 other;
DB
QY
DB
QY
DB
QY
DB

Query Match 2.5%; Score 37.8; DB 23; Length 5591;
Best Local Similarity 57.0%; Pred. No. 0.99;
Matches 69; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

QY 1329 ACATTTGGTAAAGCAGATGTGCGAGCTCTCTGAAAAATGTATTGGCGCAGGCGCTGAT 1388
DB 1358 ACATGTGGAGAGCTGATGTGGAGCGGACCTGGACCGGAGGATGCCCAACACGAGC 1417

QY 1389 GGTGCACCTGTGGAGGAGCAGGAGCTGGATGGAAAAATAAATGTGTACAGGAACGGC 1448
DB 1418 GCTGCACCTCCAGAGAACATCAACGAGCTAAAGCAAGAATCGTTGAATTGGAGTCGGC 1477

QY 1449 A 1449
DB 1478 A 1478

RESULT 27
ABL08655
ID ABL08655 standard; cDNA; 6000 BP.
XX
XX ABL08655;
XX
XX 26-MAR-2002 (first entry)
XX
XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 20447.
DE Drosophila; developmental biology; cell signalling; insecticide;
XX
XX pharmaceutical; gene; ss.
XX
XX Drosophila melanogaster.
OS
XX WO200171042-A2.
XX
XX 27-SEP-2001.
XX
XX 23-MAR-2001; 2001WO-US09231.
XX
XX 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX

PA (PEKE) PE CORP NY.
XX Venter JC, Adams M, Li PWD, Myers EW;
XX WPI; 2001-656860/75.
DR P-PSDB; ABB64552.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
XX Claim 1; SEQ ID NO 20447; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).
XX
XX The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 6000 BP; 2026 A; 1274 C; 1434 G; 1266 T; 0 other;
SQ
Query Match 2.5%; Score 37.8; DB 23; Length 6000;
Best Local Similarity 57.0%; Pred. No. 1;
Matches 69; Conservative 0; Mismatches 52; Indels 0; Gaps 0;
QY 1329 ACATTTGGTAAGCACATGGTCAGCGCTCTCTCGAAATGATTTGGCGGCGCTGAT 1388
DB 1767 ACATGTGGAGAGCTGATGGTGAGCGCGACCTGGACCGGAGGATGCCAGAACCGGC 1826
QY 1389 GGTGCACCTGTGGAGGAGGAGGACCTGGATGGAATAAATGCTGTACAGGAACGCG 1448
DB 1827 GCTGCAGCTGCAGAGAACATCAACGAGCTAAAGCAAGAACTGTTGAATGGAGTCGGC 1886
QY 1449 A 1449
DB 1887 A 1887
RESULT 28
AAAX20256
ID AAX20256 standard; DNA; 14752 BP.
XX
XX AAX20256;
AC
XX
XX 04-MAY-1999 (first entry)
DT
XX
XX Borrelia burgdorferi polynucleotide sequence #9.
DE
XX
XX Borrelia burgdorferi; spirochete; bacterium; pathogen; Lyme disease;
KW epidemic relapsing fever; endemic relapsing fever; Lyme borreliosis;
KW infection; diagnosis; characterisation; detection; ds.
XX
XX Borrelia burgdorferi.
OS
XX
XX WO9858943-A1.
PN
XX
XX 30-DEC-1998.
PD
XX
XX 18-JUN-1998; 98WO-US12764.
PF
XX
XX 03-SEP-1997; 97US-0057483.
PR
XX 20-JUN-1997; 97US-0050359.
PR
XX 22-JUL-1997; 97US-0053344.
PR
XX 22-JUL-1997; 97US-0053377.
PR
XX (HUMA-) HUMAN GENOME SCI INC.
PA
XX (MEDI-) MEDIMUNE INC.

XX Clayton R, Dougherty BA, Fraser C, Lathigra R, Smith HO;
PI White OR;
XX WPI; 1999-081217/07.
XX
XX New isolated Borrelia burgdorferi nucleic acids - used to develop
PT products for the detection, diagnosis, characterisation, prevention
PT and therapy of infections, particularly Lyme disease
XX
XX Claim 1; Page 893-901; 1128pp; English.
XX
XX AAX20248 to AAX20402 represent polynucleotide sequences isolated from
CC Borrelia burgdorferi (Bb). Products derived from Bb can be used for
CC the detection, diagnosis, characterisation, prevention and therapy of
CC Bb infections, e.g. Lyme disease. They can also be used for the
CC production of biosynthetic products, e.g. enzymes. Borrelia belongs
CC to a family of motile, spiral-shaped bacteria called Spirochetes.
CC Spirochetes are pathogenic in humans and Borrelia causes epidemic and
CC endemic relapsing fever, and Lyme borreliosis, more commonly known as
CC Lyme disease.
XX
XX Sequence 14752 BP; 5196 A; 1781 C; 1634 G; 6139 T; 2 other;
SQ
Query Match 2.5%; Score 36.6; DB 20; Length 14752;
Best Local Similarity 57.4%; Pred. No. 4.5;
Matches 66; Conservative 0; Mismatches 49; Indels 0; Gaps 0;
QY 426 TAGTATTTTGTCTGTAGTATCTATCCAGCAATAGGTATATCTCTGTCATCAATAA 485
DB 8104 TATCTTTTATTTCTTTTATTAATGCTTCTGTAATAATTATAAGCGGTGTTTCCAAAGA 8163
QY 486 GTTGACCTTTGTATACACATCGGAATTTCCCTTAATCCGGAGCTATTCGTATGA 540
DB 8164 TTTTCTTTTGTATATAAATGCAAGCTTCGCTTATCTTTGCTTAATGTTTTA 8218
RESULT 29
AAAF07750
ID AAF07750 standard; cDNA; 801 BP.
XX
XX AAF07750;
AC
XX
XX 13-MAR-2001 (first entry)
DT
XX
XX Fusarium venenatum EST SEQ ID NO:273.
DE
XX
XX Multiple gene expression; filamentous fungal cell; EST;
KW expressed sequence tag; Fusarium venenatum; Aspergillus niger;
KW Aspergillus oryzae; Trichoderma reesei; identification; recombination;
KW culture condition; environmental stress; spore morphogenesis;
KW metabolic pathway engineering; catabolic pathway engineering; ss.
XX
XX Fusarium venenatum.
OS
XX
XX WO200056762-A2.
PN
XX
XX 28-SEP-2000.
PD
XX
XX 22-MAR-2000; 2000WO-US07781.
PF
XX
XX 22-MAR-1999; 99US-0273623.
PR
XX
XX (NOVO) NOVO NORDISK BIOTECH INC.
PA
XX (NOVO) NOVO NORDISK AS.
PA
XX Berka RM, Rey MW, Shuster JR, Kauppinen S, Clausen IG, Olsen PB;
PI WPI; 2000-594572/56.
DR
XX
XX Monitoring differential expression of genes in filamentous fungal cells
PT uses fluorescence-labeled nucleic acids isolated from the cells and a
PT substrate of expressed sequence tags -

[illegible]

XX	13-FEB-2002	(first entry)
DT		
XX	DNA encoding novel human diagnostic protein #25942.	
XX		
XX	Human; chromosome mapping; gene mapping; gene therapy; forensic;	
KW	food supplement; medical imaging; diagnostic; genetic disorder; ss.	
KW		
XX	Homo sapiens.	
OS		
XX	WO200175067-A2.	
PN		
XX	11-OCT-2001.	
PD		
XX	30-MAR-2001; 2001WO-US08631.	
XX		
XX	31-MAR-2000; 2000US-0540217.	
PR		
XX	23-AUG-2000; 2000US-0649167.	
PR		
XX	(HYSE-) HYSEQ INC.	
PA		
XX	Dmanac RT, Liu C, Tang YT;	
PI		
XX	WPI; 2001-639362/73.	
DR	P-PSDB; ABG25951.	
XX		
XX	New isolated polynucleotide and encoded polypeptides, useful in	
PT	diagnostics, forensics, gene mapping, identification of mutations	
PT	responsible for genetic disorders or other traits and to assess	
PT	biodiversity -	
XX		
XX	Claim 1; SEQ ID NO 25942; 103pp; English.	
XX		
CC	The invention relates to isolated polynucleotide (I) and	
CC	polypeptide (II) sequences. (I) is useful as hybridisation probes,	
CC	polymerase chain reaction (PCR) primers, oligomers, and for chromosome	
CC	and gene mapping, and in recombinant production of (II). The	
CC	polynucleotides are also used in diagnostics as expressed sequence tags	
CC	for identifying expressed genes. (I) is useful in gene therapy techniques	
CC	to restore normal activity of (II) or to treat disease states involving	
CC	(II). (II) is useful for generating antibodies against it, detecting or	
CC	quantitating a polypeptide in tissue, as molecular weight markers and as	
CC	a food supplement. (II) and its binding partners are useful in medical	
CC	imaging of sites expressing (II). (I) and (II) are useful for treating	
CC	disorders involving aberrant protein expression or biological activity.	
CC	The polypeptide and polynucleotide sequences have applications in	
CC	diagnostics, forensics, gene mapping, identification of mutations	
CC	responsible for genetic disorders or other traits to assess biodiversity	
CC	and to produce other types of data and products dependent on DNA and	
CC	amino acid sequences. AAS64197-AA594564 represent novel human	
CC	diagnostic coding sequences of the invention.	
CC	Note: The sequence data for this patent did not appear in the printed	
CC	specification, but was obtained in electronic format directly from WIPO	
CC	at ftp.wipo.int/pub/published_pct_sequences.	
XX		
SQ	Sequence 6240 BP; 1491 A; 1747 C; 1604 G; 1398 T; 0 other;	
	Query Match 2.3%; Score 34.8; DB 23; Length 6240;	
	Best Local Similarity 49.2%; Pred. No. 10;	
	Matches 121; Conservative 0; Mismatches 122; Indels 3; Gaps 1;	
QY	797 TGGCGATTATGGTCAATTATGTTCCTTTATTCGTATGGCTTGCGACGTGCGGAAAC 856	
Db	1471 TGGCGATTCTGCTCTCTCTAGTAGAGCTGTGATCGGTGGCCTGGGCATCTGTTTAC 1412	
QY	857 ATACAGCATATATGGCGGGAGGCCAGTGTGTCAGAACGTTTGAACCGCT 916	
Db	1411 CTTCCTGGTGGCGAACAACGGGTGGTCCCAACGGTCTGGCATTAATGCCGA 1352	
QY	917 GAACAGCTGCCGGATAAATTAATTCGTGATTAACCCGTCGATTCGTGGCCAGTCAA 976	
Db	1351 GCGCAGCTGGGATGTGAACGCCGACGCCCTTCGTCTGCTGCTGCGAGAAAATCCA 1292.	

PR 06-SEP-2000; 2000US-0230438.
 PR 08-SEP-2000; 2000US-0231242.
 PR 08-SEP-2000; 2000US-0231243.
 PR 08-SEP-2000; 2000US-0231244.
 PR 08-SEP-2000; 2000US-0231413.
 PR 08-SEP-2000; 2000US-0231414.
 PR 08-SEP-2000; 2000US-0232080.
 PR 08-SEP-2000; 2000US-0232081.
 PR 12-SEP-2000; 2000US-0231968.
 PR 14-SEP-2000; 2000US-0232397.
 PR 14-SEP-2000; 2000US-0232398.
 PR 14-SEP-2000; 2000US-0232399.
 PR 14-SEP-2000; 2000US-0232400.
 PR 14-SEP-2000; 2000US-0232401.
 PR 14-SEP-2000; 2000US-0233663.
 PR 14-SEP-2000; 2000US-0233664.
 PR 14-SEP-2000; 2000US-0233665.
 PR 21-SEP-2000; 2000US-0234223.
 PR 21-SEP-2000; 2000US-0234274.
 PR 25-SEP-2000; 2000US-0234997.
 PR 25-SEP-2000; 2000US-0234998.
 PR 26-SEP-2000; 2000US-0235484.
 PR 27-SEP-2000; 2000US-0235834.
 PR 27-SEP-2000; 2000US-0235836.
 PR 29-SEP-2000; 2000US-0236327.
 PR 29-SEP-2000; 2000US-0236367.
 PR 29-SEP-2000; 2000US-0236368.
 PR 29-SEP-2000; 2000US-0236369.
 PR 29-SEP-2000; 2000US-0236370.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 02-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239935.
 PR 13-OCT-2000; 2000US-0239937.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241221.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 20-OCT-2000; 2000US-0241826.
 PR 01-NOV-2000; 2000US-0244617.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.

PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 11-DEC-2000; 2000US-0251990.
 PR 05-JAN-2001; 2001US-0259678.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 XX Rosen CA, Barash SC, Ruben SM;
 PI
 XX WPI; 2001-483426/52.
 DR P-PSDB; AAM82488.
 XX
 XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
 PT useful for preventing, diagnosing and/or treating cancers and
 PT metastasis -
 XX
 PS Claim 1; SEQ ID NO 329; 3071pp + Sequence Listing; English.
 XX
 CC AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
 CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
 CC activity, and can be used in gene therapy and vaccine production. (I)
 CC proteins and polynucleotides may be used in the prevention, diagnosis and
 CC treatment of diseases associated with inappropriate (I) expression. For
 CC example, they may be used to treat disorders associated with decreased
 CC expression by rectifying mutations or deletions in a patient's genome
 CC that affect the activity of (I) by expressing inactive proteins or to
 CC supplement the patients own production of (I). Additionally, (I)
 CC polynucleotides may be used to produce the secreted (I), by inserting
 CC the nucleic acids into a host cell and culturing the cell to express the
 CC protein. (I) proteins and polynucleotides may be used to prevent,
 CC diagnose and treat immune/hematopoietic-related diseases, especially
 CC cancers and cancer metastases of hematopoietic-derived cells. AAK64703
 CC to AAK87694 represent human immune/hematopoietic antigen genomic
 CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
 CC represent sequences used in the exemplification of the present invention.
 XX
 SQ Sequence 526 BP; 167 A; 97 C; 102 G; 159 T; 1 other;

Query Match 2.3%; Score 34.6; DB 22; Length 526;
 Best Local Similarity 52.9%; Pred.No. 2.5;
 Matches 73; Conservative 0; Mismatches 65; Indels 0; Gaps 0;
 QY 528 GCTATTGCTATGATAAAAAAAGCTCTTCCTGTTCTGATCTTCCTGGCGTATCGGGAGC 587
 Db 367 GATATAATACATAAATAATCACTCTTCTATATACATTTCCATGCTTTATGACAAAGTG 426
 QY 588 TTTTCTACCGCTGAGCGCTGATAAAAAGAGACTCAAAATTTCTACTATCCAGAAACA 647
 Db 427 TGTGTAAGTNCCTATACTACTACACAGTTGAGACITAAATTTCTACTATCTCAAAAG 486
 QY 648 CTGGATTTAACTCCTCTG 665
 Db 487 TTTCTTAGCCACTTCAG 504

RESULT 35
 ABL70580/c

ABL70580 standard; DNA; 6065 BP.

AC ABL70580;

DT 01-JUL-2002 (first entry)

DE Chemically treated cell signalling DNA sequence complementary to#235.

DE Cell signalling; cytosine methylation; cell signalling disease;

KW cancer; tumour; cytostatic; ds.

XX Unidentified.

OS

XX

PN WO200202807-A2.

PN

PD 10-JAN-2002.

PD

XX

XX

PF 29-JUN-2001; 2001WO-EP07471.

PF

XX

XX

PR 30-JUN-2000; 2000DE-1032529.

PR

PR 01-SEP-2000; 2000DE-1043826.

PR

XX

PA (EPIG-) EPIGENOMICS AG.

PA

XX

XX

PI Olek A, Piepenbrock C, Berlin K;

PI

XX

XX

DR WPI; 2002-154758/20.

DR

XX

XX

PT Nucleic acid, useful for diagnosis and therapy of diseases associated

PT with cell signalling e.g. cancer, comprises chemically modified genomic

PT sequences of genes associated with cell signalling -

XX

XX

PS Claim 1; SEQ ID NO 470; 24bp+sequence listing; English.

PS

XX

XX

CC The invention relates to a nucleic acid comprising a sequence of at least

CC 18 bases of a segment of chemically pretreated DNA of genes associated

CC with cell signalling. The activity of the modified sequences of the

CC invention may be described as cytostatic. The object of the invention is

CC to provide the chemically modified DNA of genes associated with cell

CC signalling, as well as oligonucleotides and/or PNA-oligomers for

CC detecting cytosine methylations, as well as a method which is

CC particularly suitable for the diagnosis and/or therapy of genetic and

CC epigenetic parameters of genes associated with cell signalling. The

CC chemically modified DNA provided by the invention is useful for diagnosis

CC and therapy of diseases such as solid tumours and cancer. The sequences

CC given in records ABL7011-ABL70626 represent chemically pre-treated

CC genomic DNA's of genes associated with cell signalling.

CC Note: The sequence data for this patent is not represented in the printed

CC specification, but is based on sequence information supplied by the

CC European Patent Office.

XX

XX

SQ Sequence 6065 BP; 1602 A; 178 C; 1473 G; 2812 T; 0 other;

SQ

Query Match 2.3%; Score 34.6; DB 24; Length 6065;

Best Local Similarity 58.1%; Pred. No. 12;

Matches 61; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 445 TATCTATCCGACGAATAGTATATCTGTTGCATCAATAAGTTGACTTTGTATACAC 504

DB TTTTCTTCTAAAAATACTAATCCCTTAAAAATAAAATAAATTAACCTTAAACAACAC 3094

QY 505 ATCGGAATTCCTTAAATCCGGAGCTATTCGTATGATAAAAAAA 549

DB TTTTCTTCTAAAAATACTAATCCCTTAAAAATAAAATAAATTAACCTTAAACAACAC 3094

QY 505 ATCGGAATTCCTTAAATCCGGAGCTATTCGTATGATAAAAAAA 549

DB TTTTCTTCTAAAAATACTAATCCCTTAAAAATAAAATAAATTAACCTTAAACAACAC 3094

RESULT 36

AAS61261/c

ID AAS61261 standard; DNA; 6065 BP.

XX

XX

AC AAS61261;

XX

XX

DT 29-JAN-2002 (first entry)

XX Human gene regulation-associated gene oligonucleotide #216.

DE

XX

KW Human; Gene regulation-associated gene; severe combined immunodeficiency;

KW cardiac damage; inflammatory response; Haemophilia; Werner syndrome;

KW asthma; HDR syndrome; congenital heart defect; Saethre-Chotzen syndrome;

KW renal disease; preclampsia; cardiac allograft vascular disease;

KW colorectal cancer; thyroid cancer; oesophageal cancer; ds; tumour;

KW immunostimulant; cardiant; antiinflammatory; coagulant; antiasthmatic;

KW nephrotropic; gynecological; anti-tumour; immunosuppressive; cytostatic.

OS

XX Homo sapiens.

XX

PN WO200177375-A2.

PN

PD 18-OCT-2001.

PD

XX

XX

PF 06-APR-2001; 2001WO-EP03968.

PF

XX

XX

PR 06-APR-2000; 2000DE-1019058.

PR

PR 07-APR-2000; 2000DE-1019173.

PR

PR 30-JUN-2000; 2000DE-1032529.

PR

PR 01-SEP-2000; 2000DE-1043826.

PR

XX

PA (EPIG-) EPIGENOMICS AG.

PA

XX

XX

PI Olek A, Piepenbrock C, Berlin K;

PI

XX

XX

DR WPI; 2002-017470/02.

DR

XX

XX

PT New nucleic acid sequences from chemically modified genes associated

PT with gene regulation, useful for analysing cytosine methylations for

PT diagnosis and therapy of diseases e.g. severe combined immunodeficiency

PT disease -

XX

XX

PS Claim 1; SEQ ID NO 222; 26pp; English.

PS

XX

XX

CC The invention relates to 224 nucleic acid sequences comprising at least

CC 18 bases of a chemically pretreated gene associated with gene regulation

CC selected from 43 known genes (or complementary sequences). The

CC chemical pretreatment converts cytosine bases unmethylated at the

CC 5-position to uracil or another base with hybridisation behaviour

CC dissimilar to cytosine, to enable analysis of cytosine methylations.

CC The DNA sequences, oligomers (or sets/arrays) and method are

CC useful in the diagnosis of diseases (or predisposition to diseases)

CC associated with gene regulation and in therapy of such diseases, by

CC enabling analysis of the cytosine methylation patterns of such genes,

CC kits are provided. They are especially useful in diagnosis

CC and therapy of e.g. severe combined immunodeficiency disease, cardiac

CC disorders, haemophilia, solid tumours and cancer, Werner syndrome,

CC asthma, HDR syndrome, Saethre-Chotzen syndrome, renal disease,

CC preclampsia, graft versus-host disease. The present sequence is a

CC sequence included in the sequence data for this specification and is

CC associated with the human gene regulation-associated genes.

CC Note: The sequence data for this patent did not form part

CC of the printed specification, but was obtained in electronic

CC format directly from WIPO at

CC ftp.wipo.int/pub/published_pct_sequences

XX

XX

SQ Sequence 6065 BP; 1602 A; 178 C; 1473 G; 2812 T; 0 other;

SQ

Query Match 2.3%; Score 34.6; DB 24; Length 6065;

Best Local Similarity 58.1%; Pred. No. 12;

Matches 61; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 445 TATCTATCCGACGAATAGTATATCTGTTGCATCAATAAGTTGACTTTGTATACAC 504

DB TTTTCTTCTAAAAATACTAATCCCTTAAAAATAAAATAAATTAACCTTAAACAACAC 3094

QY 505 ATCGGAATTCCTTAAATCCGGAGCTATTCGTATGATAAAAAAA 549

DB TTTTCTTCTAAAAATACTAATCCCTTAAAAATAAAATAAATTAACCTTAAACAACAC 3094

QY 524 CGGACGCTATTCGATGATAAAAAAACAATCTTCTGTTGATTTCTTCTGGCGCTATCGGG 583
DB 354 CTGCGCTGTTGCGATGCGTCAATAAATTCCTTGAGTCTTGCCCTCAAGACTCTCGTGG 295
QY 584 GAGCTTTTCTACGCTGAGCGCTGATGATAAAAAAGAGACT 623
DB 294 GCCTGTTCCACCGAGACACAGACAGACAGCCGAGTGTCT 255

RESULT 39
AAZ15175/C
ID AAZ15175 standard; cDNA; 800 BP.
AC AAZ15175;
XX
DT 12-OCT-1999 (first entry)
XX
DE Human gene expression product cDNA sequence SEQ ID NO:2644.
XX
KW Human; gene; gene expression product; diagnosis; therapy; probe;
KW detection; mapping; tissue typing; profiling; forensic; cancer;
KW genetic analysis; colorectal cancer; breast cancer; lung cancer; ss.
XX
OS Homo sapiens.
XX
PN WO9938972-A2.
XX
PD 05-AUG-1999.
XX
PF 28-JAN-1999; 99WO-US01619.
XX
PR 03-APR-1998; 98US-0080666.
PR 28-JAN-1998; 98US-0072910.
PR 24-FEB-1998; 98US-0075954.
PR 31-MAR-1998; 98US-0080114.
PR 03-APR-1998; 98US-0080515.
XX
(CHIR) CHIRON CORP.
PA (HYSE-) HYSEQ INC.
XX
PI Crkvenjakov R, Dickson M, Drmanac R, Drmanac S;
PI Escobedo J, Garcia PD, Garcia V, Glese K, Innis MA;
PI Jones WL, Kassam A, Kennedy GC, Kita D, Labat I;
PI Lamson G, Leshkowitz D, Pot D, Randazzo F, Reinhard C;
PI Stache-Crain B, Sudduth-Klinger J, Williams LT;
XX
WPI; 1999-494092/41.
XX
XX Novel human genes and their expression products which are
PT differentially expressed in different cell types
XX
PS Claim 1; Page 1286; 2479pp; English.

The present invention describes a library of human polynucleotides comprising the sequences given in AAZ12532 to AAZ1779. Also described is a method of detecting differentially expressed genes correlated with the cancerous state of a mammalian cell, comprising detecting at least one differentially expressed gene product in a test sample from a cell suspected of being cancerous, where the gene product is encoded by one of the 5248 polynucleotide sequences given in AAZ12532 to AAZ1779. The polynucleotides can be used as a source of primers and probes, which can be used for a variety of purpose, e.g. detection of expression levels, mapping, tissue typing or profiling, forensics, genetic analysis and detection of polymorphisms. Polypeptides encoded by the polynucleotides can be used for raising antibodies for experimental, diagnostic and therapeutic purposes. The polynucleotides may also be used to construct arrays for diagnostics (which may be used to determine function of an encoded protein); and to detect differences in expression levels between two cells (e.g. to identify abnormal or diseased tissue in a human, to identify a genetic predisposition or susceptibility to a disease such as cancer). The polynucleotides of the invention are especially used in the diagnosis, prognosis and management of colorectal cancer, breast cancer, and lung cancer. The polynucleotides can also be used to screen for

CC peptide analogues and antagonists.
XX
SQ Sequence 800 BP; 255 A; 153 C; 158 G; 209 T; 25 other;
Query Match 2.3%; Score 34.2; DB 20; Length 800;
Best Local Similarity 53.4%; Pred. No. 4.4;
Matches 63; Conservative 0; Mismatches 55; Indels 0; Gaps 0;
QY 611 TAAAAAGAGAGACTCAAAATTTCTACTATCCAGAAACACTGGATTAACTCTCTGAGATT 670
DB 793 TTAACCTAGCCAGGTTTCCCTGTCACCNAGACTGGAGTTAAGAACCTGANAAN 734
QY 671 ACACAGCCCTGAATCAATCCCTGGGGGCTGATTTGATATGCCACAGATTTCAA 728
DB 733 AGCCCACTGTCCTCAAAATTTCTGGGGCTCAAGTGATCTCTCTGCCTCAACNTCTAAA 676

RESULT 40
AAS49611
ID AAS49611 standard; DNA; 893 BP.
AC AAS49611;
XX
DT 13-FEB-2002 (first entry)
XX
DE Staphylococcus aureus cellular proliferation inhibitory sequence #835.
XX
KW Antisense; ss; prokaryotic cellular proliferation;
KW antibiotic; antibacterial; drug design.
XX
OS Staphylococcus aureus.
XX
PN WO200170955-A2.
XX
PD 27-SEP-2001.
XX
PF 21-MAR-2001; 2001WO-US09180.
XX
PR 21-MAR-2000; 2000US-191078P.
PR 23-MAY-2000; 2000US-206848P.
PR 26-MAY-2000; 2000US-207727P.
PR 23-OCT-2000; 2000US-242578P.
PR 27-NOV-2000; 2000US-253625P.
PR 22-DEC-2000; 2000US-257931P.
PR 16-FEB-2001; 2001US-269308P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI Yamamoto RT, Xu HH;
XX
WPI; 2001-611495/70.
XX
XX New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids -
XX
PS Claim 1; Seq ID No 2188; 511pp; English.
XX
CC The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the
CC genes, their use in the discovery of novel antibiotics, the essential
CC genes themselves and the encoded proteins. The prokaryotes used are
CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
CC invention is also useful for the identification of potential new targets
CC for antibiotic development. The antisense nucleic acids can also be used
CC to identify proteins used in proliferation, to express these proteins,
CC and to obtain antibodies capable of binding to the expressed proteins.
CC The proteins can be used to screen compounds in rational drug discovery
CC programmes. The antisense nucleic acid sequence is also useful to screen
CC for homologous nucleic acids which are required for cell proliferation in
CC a wide variety of organisms. The present sequence is an antisense
CC oligonucleotide of the invention.

CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 893 BP; 273 A; 161 C; 119 G; 340 T; 0 other;

Query Match 2.3%; Score 34.2; DB 23; Length 893;
Best Local Similarity 49.7%; Pred. No. 4.7;
Matches 87; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

QY 483 AAGTTGACATTTGTATACACATGCGAATTTCCCTTAATCCGGAGCTATTCGTATGATA 542
DB 30 AAGTTCAAAATGCTTCTGCATCTTGGAAATGCTTTTATCAGCAACTCATCTAATA 89
QY 543 AAAAAACATCTTCTGTTCTGATCTCTGCGCTATTCGGGGAGCTTTTCTACCGCTGTA 602
DB 90 ATAATTTTAAATTTATCTGTTCTCTAGTAGCACTTGTCCAGGATTAAGAAACTATACG 149
QY 603 GCCGCTGATAAAAAAGAGACTCAAAATTTCTACTATCCAGAACACATGGATTAA 657
DB 150 TCGGCTCTAGTACAAATGATACCAAAATGATATCATCGATAAGTCTTCTGTTTAA 204

RESULT 41
AAS55352/C
ID AAS55352 standard; DNA; 927 BP.

AC AAS55352;
DT 13-FEB-2002 (first entry)

DE Staphylococcus aureus DNA for cellular proliferation protein #1664.
XX Antisense; ds; prokaryotic cellular proliferation gene;
KW antibiotic; antibacterial; drug design.

OS Staphylococcus aureus.

PN WO200170955-A2.

PD 27-SEP-2001.

PF 21-MAR-2001; 2001WO-US09180.

PR 21-MAR-2000; 2000US-191078P.

PR 23-MAY-2000; 2000US-206848P.

PR 26-MAY-2000; 2000US-207727P.

PR 23-OCT-2000; 2000US-242578P.

PR 27-NOV-2000; 2000US-253625P.

PR 22-DEC-2000; 2000US-257931P.

PR 16-FEB-2001; 2001US-269308P.

XX (ELIT-) ELITRA PHARM INC.

XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;

XX Yamamoto RT, Xu HH;

XX WPI; 2001-611495/70.

XX P-PSDB; AAU37493.

XX New polynucleotides for the identification and development of

XX antibiotics, comprise sequences of antisense nucleic acids -

XX Claim 27; Seq ID No 8989; 51lpp; English.

XX The invention relates to antisense inhibitors of genes essential to

XX prokaryotic cellular proliferation, their use in identifying the

XX genes, their use in the discovery of novel antibiotics, the essential

CC for antibiotic development. The antisense nucleic acids can also be used
CC to identify proteins used in proliferation, to express these proteins,
CC and to obtain antibodies capable of binding to the expressed proteins.
CC The proteins can be used to screen compounds in rational drug discovery
CC programmes. The antisense nucleic acid sequence is also useful to screen
CC for homologous nucleic acids which are required for cell proliferation in
CC a wide variety of organisms. The present sequence encodes an
CC essential prokaryotic cellular proliferation protein.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 927 BP; 347 A; 116 C; 169 G; 295 T; 0 other;

Query Match 2.3%; Score 34.2; DB 23; Length 927;
Best Local Similarity 49.7%; Pred. No. 4.8;
Matches 87; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

QY 483 AAGTTGACATTTGTATACACATGCGAATTTCCCTTAATCCGGAGCTATTCGTATGATA 542
DB 715 AAGTTCAAAATGCTTCTGCATCTTGGAAATGCTTTTATCAGCAACTCATCTAATA 656
QY 543 AAAAAACATCTTCTGTTCTGATCTCTGCGCTATTCGGGGAGCTTTTCTACCGCTGTA 602
DB 655 ATAATTTTAAATTTATCTGTTCTCTAGTAGCACTTGTCCAGGATTAAGAAACTATACG 596
QY 603 GCCGCTGATAAAAAAGAGACTCAAAATTTCTACTATCCAGAACACATGGATTAA 657
DB 595 TCGGCTCTAGTACAAATGATACCAAAATGATATCATCGATAAGTCTTCTGTTTAA 541

RESULT 42
AAS52250/C
ID AAS52250 standard; DNA; 1164 BP.

AC AAS52250;
DT 13-FEB-2002 (first entry)

DE Staphylococcus aureus DNA for cellular proliferation protein #667.
XX Antisense; ds; prokaryotic cellular proliferation gene;
KW antibiotic; antibacterial; drug design.

OS Staphylococcus aureus.

PN WO200170955-A2.

PD 27-SEP-2001.

PF 21-MAR-2001; 2001WO-US09180.

PR 21-MAR-2000; 2000US-191078P.

PR 23-MAY-2000; 2000US-206848P.

PR 26-MAY-2000; 2000US-207727P.

PR 23-OCT-2000; 2000US-242578P.

PR 27-NOV-2000; 2000US-253625P.

PR 22-DEC-2000; 2000US-257931P.

PR 16-FEB-2001; 2001US-269308P.

XX (ELIT-) ELITRA PHARM INC.

XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;

XX Yamamoto RT, Xu HH;

XX WPI; 2001-611495/70.

XX P-PSDB; AAU34391.

XX New polynucleotides for the identification and development of

XX antibiotics, comprise sequences of antisense nucleic acids -

XX Claim 27; Seq ID No 4832; 51lpp; English.

XX The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the
CC genes, their use in the discovery of novel antibiotics, the essential
CC themselves and the encoded proteins. The prokaryotes used are
CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
CC invention is also useful for the identification of potential new targets
CC for antibiotic development. The antisense nucleic acids can also be used
CC to identify proteins used in proliferation, to express these proteins,
CC and to obtain antibodies capable of binding to the expressed proteins.
CC The proteins can be used to screen compounds in rational drug discovery
CC programmes. The antisense nucleic acid sequence is also useful to screen
CC for homologous nucleic acids which are required for cell proliferation in
CC a wide variety of organisms. The present sequence encodes an
CC essential prokaryotic cellular proliferation protein.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 1164 BP; 426 A; 164 C; 203 G; 371 T; 0 other;

Query Match 2.3%; Score 34.2; DB 23; Length 1164;
Best Local Similarity 49.7%; Pred. No. 5.6;
Matches 87; Conservative 0; Mismatches 88; Indels 0; Gaps 0;
QY 483 AAGTTGACTTTTGTATACACATGCGAATTTCCCTTAATCCGGAGCTATTTCGTATGATA 542
DB 1021 AAGTTCAAGCTTCGTCATCTTGGAAATGCGCTTTTATCAGCACTCATCTAATA 962
QY 543 AAAAAAATCTTCCTGTTCTGATTCTCTGCGCTATCGGGAGCTTTTCTACCGCTGTA 602
DB 961 ATAATTTTAAATTTATCTGTTCTCAGTACGACTTGTACGATTTGAAAACTATACG 902
QY 603 GCGCTGTATAAAGAGACTCAAAATTTCTACTATCCAGAACACATGGATTAA 657
DB 901 TCGGCTCTAGTACAAATGTATACCAATGTATCATCGATTAAGTCCTTCGTTTAA 847

RESULT 43
AAV74876
ID AAV74876 standard; DNA; 2554 BP.

XX AAV74876;

XX 16-MAR-1999 (first entry)

XX Staphylococcus aureus contig SEQ ID #565.

XX Computer readable medium; vaccine; S.aureus infection; Immunodetection;
KW cellulitis; eyelid infection; food poisoning; osteomyelitis; therapy;
KW skin infection; surgical wound infection; scalded skin syndrome;
KW toxic shock syndrome; ds.

XX Staphylococcus aureus.

XX Key Location/Qualifiers
FH misc_feature 1561..1620
FT /tag= a
FT /note= "these bases represent a line of missing text in
FT the sequence listing in the specification. They
FT are included to maintain the nucleotide numbering
FT given in the specification for this DNA sequence"

XX EP786519-A2.

XX 30-JUL-1997.

XX 07-JAN-1997; 97EP-0100117.

XX 05-JAN-1996; 96US-0009861.

PA (HUMA-) HUMAN GENOME SCI INC.
XX Barash SC, Choi GH, Dillon PJ, Fannon MR, Kunsch CA;
PI Rosen CA;
XX WPI; 1997-374922/35.
XX Polynucleotide(s) and proteins derived from Staphylococcus aureus
PT stored on computer readable medium and used in the production of
PT anti-S.aureus vaccines

XX Claim 1; Page 1503-1504; 3271pp; English.

XX This sequence represents one of 5191 Staphylococcus aureus DNA sequences
CC of the invention. The DNA sequences are recorded on a computer readable
CC medium, preferably selected from a floppy or hard disk, random access
CC memory (RAM), read-only memory (ROM) or CD-ROM. Homology searches using
CC the S.aureus DNA sequences allows putative functions to be assigned so
CC that protein-encoding or regulatory regions of commercial, therapeutic or
CC industrial importance can be obtained. Specifically, sequences which are
CC likely to encode antigens have been identified and these polypeptides can
CC be used in a vaccine composition against S.aureus infection. The
CC polypeptides can also be used in a kit for the immunodetection of
CC S.aureus in a sample. S.aureus is implicated in numerous human diseases,
CC including cellulitis, eyelid infections, food poisoning, osteomyelitis,
CC skin and surgical wound infections, scalded skin syndrome, toxic shock
CC syndrome, etc. Organisms transformed with the DNA sequences can be used
CC for recombinant production of the polypeptides. The new DNA sequences
CC (and their fragments) are useful as primers or probes for isolating
CC homologues of any of the S.aureus DNA sequences contained on the
CC computer readable medium.

XX Sequence 2554 BP; 793 A; 433 C; 328 G; 933 T; 67 other;

Query Match 2.3%; Score 34.2; DB 18; Length 2554;
Best Local Similarity 49.7%; Pred. No. 9.1;
Matches 87; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

QY 483 AAGTTGACTTTTGTATACACATGCGAATTTCCCTTAATCCGGAGCTATTTCGTATGATA 542
DB 405 AAGTTCAAAATGCTTCGTCATCTTGGAAATGCGCTTTTATCAGCACTCATCTAATA 464
QY 543 AAAAAAATCTTCCTGTTCTGATTCTCTGCGCTATCGGGAGCTTTTCTACCGCTGTA 602
DB 465 ATAATTTTAAATTTATCTGTTCTTCAGTACGACTTGTACGATTTGAAAACTATACG 524
QY 603 GCGCTGTATAAAGAGACTCAAAATTTCTACTATCCAGAACACATGGATTAA 657
DB 525 TCGGCTCTAGTACAAATGTATACCAAAATGTATCATCGATTAAGTCCTTCGTTTAA 579

RESULT 44
AAS99435/c
ID AAS99435 standard; cDNA; 188 BP.

XX AAS99435;

XX 12-MAR-2002 (first entry)

XX Human cDNA fragment for cancer related protein 98P7C3.

XX Human; 98P6C3; ss; homeodomain protein; vaccine; cytostatic.
KW epitope; transgenic animal; immunogen; T cell; B cell; cytotoxic T cell;
KW C1u; prostate cancer; bladder cancer; kidney cancer; lung cancer;
KW breast cancer; uterine cancer; cervical cancer; stomach cancer;
KW rectal cancer; colon cancer; chromosome 4q11-q12; PCR primer; adapter;
KW suppression subtractive hybridisation; SSH.

XX Homo sapiens.

XX WO200190157-A2.

XX 29-NOV-2001.

XX PF 24-MAY-2001; 2001WO-US17495.
 XX PR 24-MAY-2000; 2000US-207138P.
 XX PA (UROC-) UROGENESYS INC.
 XX PI Challita-Eid PM, Hubert RS, Faris M, Afar DEH, Levin E;
 XX PI Mitchell SC, Jakobovits A;
 XX DR WPI; 2002-097642/13.
 XX PT New isolated 98p7C3-related homeodomain protein highly expressed in
 XX PT various cancers, useful in cancer vaccines and for generating immune
 XX PT response directed to 98p7C3 in mammal
 XX PS Example 1; Fig 1; 155pp; English.
 XX CC The invention relates to an isolated 98p7C3-related protein which is a
 XX CC homeodomain protein highly expressed in various cancers. Also include
 XX CC are polynucleotides encoding the protein or proteins 90% identical
 XX CC to 98p7C3, a pharmaceutical composition comprising the polynucleotides
 XX CC (including an expression vector comprising the 98p7C3 encoding
 XX CC polynucleotides) or a host cell transformed with the vector,
 XX CC an anti-98p7C3 antibody, a non-human transgenic animal expressing a
 XX CC 98p7C3 protein, methods of detecting the 98p7C3 protein or
 XX CC polynucleotides in a biological sample, monitoring the presence of
 XX CC cancer in an individual by detecting an elevated level of the 98p7C3
 XX CC protein or polynucleotides and a pharmaceutical composition comprising a
 XX CC modulator of 98p7C3 protein, or T cell/B cell epitopes derived
 XX CC from it, are useful in inducing an immune response (in mammal) to a
 XX CC 98p7C3 protein. Upon contact with a cytotoxic T cell (CTL) the immunogens
 XX CC induce the CTL (with its helper T cell) to kill an autologous cell
 XX CC expressing 98p7C3. The immunogen may be a nucleic acid encoding the
 XX CC protein or epitope. The antibody is useful for delivering a cytotoxic
 XX CC agent to a cell that expresses 98p7C3, by conjugating the cytotoxic
 XX CC to the antibody or its fragment that specifically binds to a 98p7C3
 XX CC epitope, and exposing the cell to the antibody-agent conjugate. The
 XX CC modulator is useful for treating a patient with a cancer that expresses
 XX CC 98p7C3 (e.g. prostate cancer, bladder cancer, kidney cancer, lung cancer,
 XX CC breast cancer, uterine cancer, cervical cancer, stomach cancer, rectal
 XX CC cancer and colon cancer), by administering to the patient a vector
 XX CC that comprises the modulator, such that the vector delivers a single
 XX CC chain monoclonal antibody coding sequence to the cancer cells and the
 XX CC encoded single chain antibody is expressed intracellularly in it.
 XX CC The gene for 98p7C3 is located on human chromosome 4q11-q12. The present
 XX CC sequence is oligonucleotide adapter or PCR primer used to isolate a cDNA
 XX CC hybridisation, SSH.
 XX SQ Sequence 188 BP; 55 A; 19 C; 15 G; 80 T; 19 other;
 Query Match 2.3%; Score 34; DB 24; Length 188;
 Best Local Similarity 57.8%; Pred. No. 2.1;
 Matches 52; Conservative 0; Mismatches 38; Indels 0; Gaps 0;
 QY 460 TAGGTATATCCGTTCATCAATAAAGTTGACTTTTGTATACAAATGCGAATTTCCCTT 519
 Db 113 TNAGTTATTTTGTNTNAATAAANTGAACCTGCTTTTGTCTCTAAATATNTTTTNAATN 54
 QY 520 ATCCGGAGCTATTCGTATGATATAAAAAA 549
 Db 53 AACCTTAACCTTTTCTTGTGGAAAAA 24
 RESULT 45
 AAH94492
 ID AAH94492 standard; cDNA; 308 BP.
 XX AC AAH94492;
 XX XX
 XX DT 05-OCT-2001 (first entry)
 XX DE Human foetal cDNA, SEQ ID NO: 520.

DE Human foetal cDNA, SEQ ID NO: 1179.
 XX Human; foetal protein; cytostatic; immunosuppressive; immunostimulant;
 KW nontropic; neuroprotective; thrombolytic; osteopathic; antiinflammatory;
 KW gene therapy; antitense therapy; cancer; immune disorder;
 KW growth disorder; osteoporosis; thrombolytic disorder;
 KW nervous system disorder; inflammation; expressed sequence tag; EST; ss.
 OS Homo sapiens.
 XX WO200155339-A2.
 XX PN 02-AUG-2001.
 XX PD 25-JAN-2001; 2001WO-US02723.
 XX PF 25-JAN-2000; 2000US-0491404.
 XX PR 15-SEP-2000; 2000US-0663870.
 XX PR 06-NOV-2000; 2000US-0707351.
 XX XX (HYSE-) HYSEQ INC.
 XX XX Yeung G, Ford JE, Boyle BJ, Arterburn MC, Drmanac RA, Tang YT;
 PI Liu C, Asundi V, Zhou P, Werhman T;
 XX WPI; 2001-465571/50.
 DR P-PSDB; AAM06817.
 XX Novel fetal proteins useful for the treatment and diagnosis of diseases
 XX associated with dysfunction of the protein e.g. cancers, immune
 XX disorders, growth disorders, thrombolytic disorders, nervous system
 XX disorders and inflammation -
 PS Example 3; Page 666; 715pp; English.
 CC The invention relates to novel foetal polypeptides encoded by
 CC polynucleotides comprising one of 477 sequences fully defined in the
 CC specification. The foetal polynucleotides and polypeptides are
 CC useful in the treatment and diagnosis of diseases such as cancers,
 CC immune disorders, growth disorders (e.g. osteoporosis), thrombolytic
 CC disorders, nervous system disorders and inflammation. The present
 CC sequence is a full length cDNA which was assembled using expressed
 CC sequence tags (ESTs) found to be expressed in human foetal tissue
 CC cDNA libraries as seeds.
 XX SQ Sequence 308 BP; 59 A; 63 C; 46 G; 140 T; 0 other;
 Query Match 2.3%; Score 34; DB 22; Length 308;
 Best Local Similarity 53.8%; Pred. No. 2.8;
 Matches 70; Conservative 0; Mismatches 60; Indels 0; Gaps 0;
 QY 425 TTAGTATTTTTCCTGTGTATCTATCCAGCAATAGGTATATCTGTGTCATCAATA 484
 Db 45 TAATTTCTTTCTTAGCGCTCTGTACTGATGAATATATATCTCTTTGGAGGTATCA 104
 QY 485 AGTTGACTTTTGTATACACATCGGAATTTCCCTTAATCCGGAGCTATTGTTATGATAA 544
 Db 105 TGTTCCTCTCTTTTTCATGTTTCTTGTGTCTTACATTAAATATCTGCACATCTGTATGA 164
 QY 545 AAAAACTCTT 554-
 Db 165 AAGTCACCT 174
 RESULT 46
 AAH93991
 ID AAH93991 standard; cDNA; 316 BP.
 XX AC AAH93991;
 XX XX
 XX DT 05-OCT-2001 (first entry)
 XX DE Human foetal cDNA, SEQ ID NO: 520.

PR 14-SEP-2000; 2000US-0232399.
 PR 14-SEP-2000; 2000US-0232400.
 PR 14-SEP-2000; 2000US-0232401.
 PR 14-SEP-2000; 2000US-0233063.
 PR 14-SEP-2000; 2000US-0233064.
 PR 14-SEP-2000; 2000US-0233065.
 PR 21-SEP-2000; 2000US-0234223.
 PR 21-SEP-2000; 2000US-0234274.
 PR 23-SEP-2000; 2000US-0234997.
 PR 25-SEP-2000; 2000US-0234998.
 PR 26-SEP-2000; 2000US-0235484.
 PR 27-SEP-2000; 2000US-0235834.
 PR 27-SEP-2000; 2000US-0235836.
 PR 29-SEP-2000; 2000US-0236327.
 PR 29-SEP-2000; 2000US-0236367.
 PR 29-SEP-2000; 2000US-0236368.
 PR 29-SEP-2000; 2000US-0236369.
 PR 29-SEP-2000; 2000US-0236370.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 02-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239935.
 PR 13-OCT-2000; 2000US-0239937.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241221.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 20-OCT-2000; 2000US-0241826.
 PR 01-NOV-2000; 2000US-0244617.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.

PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251889.
 PR 11-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX
 XX Rosen CA, Barash SC, Ruben SM;
 PI WPI; 2001-476223/51.
 XX
 XX Novel isolated prostate gland related polypeptide useful for diagnosis
 PT and treatment of disorders of prostate such as prostatodystonia,
 PT prostatosis, prostatitis, benign prostatic hypertrophy and malacoplakia
 PT
 XX
 PS Claim 1; SEQ ID No 308; 512pp; English.
 XX
 CC The invention relates to novel isolated prostate gland related nucleic
 CC acids (I) and polypeptides (II). (I) and (II) are useful for diagnosis,
 CC prognosis, prevention, and/or treatment of diseases and/or disorders of
 CC the prostate such as acute non-bacterial prostatitis, chronic non-
 CC bacterial prostatitis, acute bacterial prostatitis, prostatodystonia,
 CC prostatosis, granulomatous prostatitis, malacoplakia, benign prostatic
 CC hypertrophy or hyperplasia, and prostate neoplastic disorders, including
 CC adenocarcinomas, transitional cell carcinomas, ductal carcinomas, and
 CC squamous cell carcinomas. (I), (II) and antibody to (II) are useful for
 CC diagnosing and treating reproductive system disorders (Paget's disease),
 CC autoimmune disorders (systemic lupus erythematosus, rheumatoid
 CC arthritis), blood-related disorders (sickle cell anaemia),
 CC hyperproliferative disorders, urinary system disorders
 CC (glomerulonephritis), cardiovascular disorders (arrhythmias),
 CC respiratory disorders, musculoskeletal system disorders, neural activity
 CC and neurological disorders (Alzheimer's disease and Parkinson's disease),
 CC endocrine disorders (Addison's disease), gastrointestinal disorders
 CC (inflammatory disorders), liver disorders (biliary liver cirrhosis),
 CC pancreatic and gall bladder disorders, disorders of the large intestine,
 CC developmental and inherited disorders, diseases at the cellular level,
 CC and wound healing and epithelial cell proliferation. (I) or (II) is
 CC useful to prevent skin aging, for preventing hair loss, to maintain
 Query Match 2.3%; Score 34; DB 22; Length 636;
 Best Local Similarity 59.2%; Pred No 4.4;
 Matches 58; Conservative 0; Mismatches 40; Indels 0; Gaps 0;
 QY 662 TCTGAGATTACACAGCCCTGAATCAATCCCTGGGGGCTGATTTGATTATGCCACGAG 721
 Db 573 TCTGTGTGCTCAGGCTGCTCAAACTCCTGCTAGCTGTTGATTTTGAGCCACTAA 514
 QY 722 ATTCAACAGCTGGATATGAGGCTCTGAAAAAGATA 759
 Db 513 GTAACAACAGCTTAACATGAATGATATATGAGTCAGCTA 476
 RESULT 48
 AAL06230/C
 ID AAL06230 standard; DNA: 636 BP.
 XX
 AC AAL06230;
 XX
 DT 21-NOV-2001 (first entry)
 XX
 DE Human reproductive system related antigen DNA SEQ ID NO: 8918.
 KW Human; reproductive system related antigen; reproductive system disorder;
 KW cancer; gene therapy; ds.
 XX
 OS Homo sapiens.

XX WO200155320-A2.
XX PD 02-AUG-2001.
XX PF 17-JAN-2001; 2001WO-US01339.
XX 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.

PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249284.
PR 17-NOV-2000; 2000US-0249285.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
(HUMA-) HUMAN GENOME SCI INC.
XX PA XX

PI Rosen CA, Barash SC, Ruben SM;
 XX WPI; 2001-465570/50.
 XX Isolated nucleic acid molecule encoding a reproductive system antigen
 PT is used in preventing, treating or ameliorating a medical condition -
 XX Disclosure; SEQ ID NO 8918; 1297pp + Sequence Listing; English.
 XX The present invention provides the protein and coding sequences of a
 CC number of human reproductive system related antigens. These can be used
 CC in the prevention and treatment of reproductive system disorders,
 CC including cancer. The present sequence is a genomic sequence encoding a
 CC protein of the invention.
 XX Sequence 636 BP; 153 A; 133 C; 129 G; 221 T; 0 other;
 SQ
 Query Match 2.3%; Score 34; DB 22; Length 636;
 Best Local Similarity 59.2%; Pred. No. 4.4;
 Matches 58; Conservative 0; Mismatches 40; Indels 0; Gaps 0;
 QY 662 TCTGAGATTACAGCCCTGTAATCAATCCCTGGGGGCTGATTTTGTATTATGCCACAG 721
 Db III III III III III III III III III III III III III III III
 QY 573 TCTTGTGTTGCTCAGGCTGCTCAAACTCCTGCTAGCTGTTGATTTTGTAGCCACTAA 514
 Db -
 QY 722 ATTTCACAGCTGGATATGAGGCTCTGAAAAAGATA 759
 Db 513 GTAAACACAGCTTAACATGAAGTATATGAATGACGCTA 476

RESULT 49
 AAQ12000/C
 ID AAQ12000 standard; cDNA; 1844 BP.
 XX
 AC AAQ12000;
 XX
 DT 25-SEP-1991 (first entry)
 XX
 DE Maize nitrite reductase.
 XX
 KW Nitrite reductase; nitrate; maize; spinach; probe; NR; ss.
 XX
 OS Zea mays L.
 XX
 FH Key Location/Qualifiers
 FT CDS 9..1709
 FT /*tag= a
 FT /product= NR
 XX
 PN US4997930-A.
 XX
 PD 05-MAR-1991.
 XX
 PD 16-MAR-1989; 89US-0324154.
 XX
 PD 16-MAR-1989; 89US-0324154.
 XX
 PA (CIBA) CIBA-GEIGY CORP.
 XX
 PI Lahners KN, Rothstein SJ;
 XX
 DR WPI; 1991-206821/28.
 DR P-PSDB; AARI2517.
 XX
 XX Cloned DNA sequence encoding maize nitrite reductase - useful for
 PT determin. of NR number and study of NR mRNA regulation and nitrate
 PT assimilation.
 XX
 PS Claim 2; Page 9-10; 10pp; English.
 XX
 CC The sequence may be used to determine the number of NR genes per
 CC maize genome and to study the effect of various environmental
 CC conditions, incl. the presence of nitrate on the level of NR mRNA of

CC maize plants grown under selected conditions. The cDNA may also be
 CC used to study NR mRNA regulation in maize. The cDNA is obtcd. using
 CC lambda gt11 as the cloning vector and a spinach NR cDNA as
 CC heterologous probe. The cDNA clone (pCIB808) is 66% homologous at
 CC the nucleotide level with the spinach NR gene, however the maize DNA
 CC has a high G/C content (69.5%) as compared to that of the spinach
 CC gene (46%) and is 75% homologous at the amino acid level.
 XX
 SQ Sequence 1844 BP; 338 A; 582 C; 655 G; 269 T; 0 other;
 Query Match 2.3%; Score 34; DB 12; Length 1844;
 Best Local Similarity 56.1%; Pred. No. 8.6;
 Matches 64; Conservative 0; Mismatches 50; Indels 0; Gaps 0;
 QY 169 GAAGCATCAGAGGGCGGGGAACTCCGTCGGCCGAGTGAACGTCGCCACTCCCGGCA 228
 Db III III III III III III III III III III III III III III III
 QY 229 GTACATGCCCGCGGCGCTGTATACCGGCAAGATGTCGCAAACTCCCGTCCCT 282
 Db -
 QY 524 GGTGAGCGCGACGGCGGAGCGCTCCAGGATGGCGGGACGTCGGGGAGCGT 471
 Db -
 RESULT 50
 AAK65489
 ID AAK65489 standard; DNA; 16605 BP.
 XX
 AC AAK65489;
 XX
 DT 06-NOV-2001 (first entry)
 XX
 DE Human Immune/haematopoietic antigen genomic sequence SEQ ID NO:20301.
 XX
 KW Human; Immune; haematopoietic; immune/haematopoietic antigen; cancer;
 KW cytostatic; gene therapy; vaccine; metastasis; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200157182-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 17-JAN-2001; 2001WO-US01354.
 XX
 PR 31-JAN-2000; 2000US-0179065.
 PR 04-FEB-2000; 2000US-0180628.
 PR 24-FEB-2000; 2000US-0184664.
 PR 02-MAR-2000; 2000US-0186350.
 PR 16-MAR-2000; 2000US-0189874.
 PR 17-MAR-2000; 2000US-0190076.
 PR 18-APR-2000; 2000US-0198123.
 PR 19-MAY-2000; 2000US-0205515.
 PR 07-JUN-2000; 2000US-0209467.
 PR 28-JUN-2000; 2000US-0214886.
 PR 30-JUN-2000; 2000US-0215135.
 PR 07-JUL-2000; 2000US-0216647.
 PR 11-JUL-2000; 2000US-0217487.
 PR 14-JUL-2000; 2000US-0217496.
 PR 14-JUL-2000; 2000US-0218290.
 PR 26-JUL-2000; 2000US-0220963.
 PR 26-JUL-2000; 2000US-0220964.
 PR 14-AUG-2000; 2000US-0224518.
 PR 14-AUG-2000; 2000US-0224519.
 PR 14-AUG-2000; 2000US-0225213.
 PR 14-AUG-2000; 2000US-0225214.
 PR 14-AUG-2000; 2000US-0225266.
 PR 14-AUG-2000; 2000US-0225267.
 PR 14-AUG-2000; 2000US-0225268.
 PR 14-AUG-2000; 2000US-0225270.
 PR 14-AUG-2000; 2000US-0225447.
 PR 14-AUG-2000; 2000US-0225757.
 PR 14-AUG-2000; 2000US-0225758.

PR	14-AUG-2000;	2000US-0225759;
PR	18-AUG-2000;	2000US-0226279;
PR	22-AUG-2000;	2000US-0226681;
PR	22-AUG-2000;	2000US-0226868;
PR	23-AUG-2000;	2000US-0227182;
PR	23-AUG-2000;	2000US-0227009;
PR	30-AUG-2000;	2000US-0228924;
PR	01-SEP-2000;	2000US-0229287;
PR	01-SEP-2000;	2000US-0229344;
PR	01-SEP-2000;	2000US-0229343;
PR	03-SEP-2000;	2000US-0229345;
PR	05-SEP-2000;	2000US-0229509;
PR	06-SEP-2000;	2000US-0229513;
PR	06-SEP-2000;	2000US-0230437;
PR	06-SEP-2000;	2000US-0230438;
PR	08-SEP-2000;	2000US-0231242;
PR	08-SEP-2000;	2000US-0231243;
PR	08-SEP-2000;	2000US-0231244;
PR	08-SEP-2000;	2000US-0231413;
PR	08-SEP-2000;	2000US-0231414;
PR	08-SEP-2000;	2000US-0232080;
PR	08-SEP-2000;	2000US-0232081;
PR	12-SEP-2000;	2000US-0231968;
PR	14-SEP-2000;	2000US-0232397;
PR	14-SEP-2000;	2000US-0232398;
PR	14-SEP-2000;	2000US-0232399;
PR	14-SEP-2000;	2000US-0232400;
PR	14-SEP-2000;	2000US-0232401;
PR	14-SEP-2000;	2000US-0233063;
PR	14-SEP-2000;	2000US-0233064;
PR	14-SEP-2000;	2000US-0233065;
PR	21-SEP-2000;	2000US-0234223;
PR	21-SEP-2000;	2000US-0234274;
PR	25-SEP-2000;	2000US-0234997;
PR	26-SEP-2000;	2000US-0234998;
PR	27-SEP-2000;	2000US-0235834;
PR	27-SEP-2000;	2000US-0235836;
PR	29-SEP-2000;	2000US-0236327;
PR	29-SEP-2000;	2000US-0236367;
PR	29-SEP-2000;	2000US-0236368;
PR	29-SEP-2000;	2000US-0236369;
PR	02-OCT-2000;	2000US-0236370;
PR	02-OCT-2000;	2000US-0236802;
PR	02-OCT-2000;	2000US-0237037;
PR	02-OCT-2000;	2000US-0237038;
PR	02-OCT-2000;	2000US-0237039;
PR	02-OCT-2000;	2000US-0237040;
PR	13-OCT-2000;	2000US-0239935;
PR	13-OCT-2000;	2000US-0239937;
PR	20-OCT-2000;	2000US-0240960;
PR	20-OCT-2000;	2000US-0241221;
PR	20-OCT-2000;	2000US-0241785;
PR	20-OCT-2000;	2000US-0241786;
PR	20-OCT-2000;	2000US-0241787;
PR	20-OCT-2000;	2000US-0241808;
PR	20-OCT-2000;	2000US-0241809;
PR	01-NOV-2000;	2000US-02441826;
PR	01-NOV-2000;	2000US-02446177;
PR	08-NOV-2000;	2000US-0246474;
PR	08-NOV-2000;	2000US-0246475;
PR	08-NOV-2000;	2000US-0246476;
PR	08-NOV-2000;	2000US-0246527;
PR	08-NOV-2000;	2000US-0246528;
PR	08-NOV-2000;	2000US-0246529;
PR	08-NOV-2000;	2000US-0246532;
PR	08-NOV-2000;	2000US-0246609;
PR	08-NOV-2000;	2000US-0246610;
PR	08-NOV-2000;	2000US-0246611;
PR	08-NOV-2000;	2000US-0246612;
PR	08-NOV-2000;	2000US-0246613;
PR	08-NOV-2000;	2000US-0246614;
PR	08-NOV-2000;	2000US-0246615;
PR	08-NOV-2000;	2000US-0246616;
PR	08-NOV-2000;	2000US-0246617;
PR	08-NOV-2000;	2000US-0246618;
PR	08-NOV-2000;	2000US-0246619;
PR	08-NOV-2000;	2000US-0246620;
PR	08-NOV-2000;	2000US-0246621;
PR	08-NOV-2000;	2000US-0246622;
PR	08-NOV-2000;	2000US-0246623;
PR	08-NOV-2000;	2000US-0246624;
PR	08-NOV-2000;	2000US-0246625;
PR	08-NOV-2000;	2000US-0246626;
PR	08-NOV-2000;	2000US-0246627;
PR	08-NOV-2000;	2000US-0246628;
PR	08-NOV-2000;	2000US-0246629;
PR	08-NOV-2000;	2000US-0246630;
PR	08-NOV-2000;	2000US-0246631;
PR	08-NOV-2000;	2000US-0246632;
PR	08-NOV-2000;	2000US-0246633;
PR	08-NOV-2000;	2000US-0246634;
PR	08-NOV-2000;	2000US-0246635;
PR	08-NOV-2000;	2000US-0246636;
PR	08-NOV-2000;	2000US-0246637;
PR	08-NOV-2000;	2000US-0246638;
PR	08-NOV-2000;	2000US-0246639;
PR	08-NOV-2000	

PR	08-NOV-2000;	2000US-0246611.
PR	08-NOV-2000;	2000US-0246613.
PR	17-NOV-2000;	2000US-0249207.
PR	17-NOV-2000;	2000US-0249208.
PR	17-NOV-2000;	2000US-0249209.
PR	17-NOV-2000;	2000US-0249210.
PR	17-NOV-2000;	2000US-0249211.
PR	17-NOV-2000;	2000US-0249212.
PR	17-NOV-2000;	2000US-0249213.
PR	17-NOV-2000;	2000US-0249214.
PR	17-NOV-2000;	2000US-0249215.
PR	17-NOV-2000;	2000US-0249216.
PR	17-NOV-2000;	2000US-0249217.
PR	17-NOV-2000;	2000US-0249218.
PR	17-NOV-2000;	2000US-0249244.
PR	17-NOV-2000;	2000US-0249245.
PR	17-NOV-2000;	2000US-0249264.
PR	17-NOV-2000;	2000US-0249265.
PR	17-NOV-2000;	2000US-0249297.
PR	17-NOV-2000;	2000US-0249299.
PR	17-NOV-2000;	2000US-0249300.
PR	01-DEC-2000;	2000US-0250160.
PR	01-DEC-2000;	2000US-0250391.
PR	05-DEC-2000;	2000US-0251030.
PR	05-DEC-2000;	2000US-0251988.
PR	05-DEC-2000;	2000US-0256719.
PR	06-DEC-2000;	2000US-0251479.
PR	08-DEC-2000;	2000US-0251856.
PR	08-DEC-2000;	2000US-0251868.
PR	08-DEC-2000;	2000US-0251869.
PR	08-DEC-2000;	2000US-0251989.
PR	08-DEC-2000;	2000US-0251990.
PR	11-DEC-2000;	2000US-0254097.
PR	05-JAN-2001;	2001US-0259678.
XX		
FA	(HUMA-) HUMAN GENOME SCI INC	
XX		
PI	Rosen CA, Barash SC, Ruben	
XX		
XX	WPT; 2001-483426/52.	
DR		
PT	Nucleic acids encoding human	
PT	useful for preventing, diagn	
PT	metastasis -	
XX		
PS	Disclosure; SEQ ID NO 20301;	
XX		
CC	AAK54951 to AAK64702 encode	
CC	amino acid sequences given i	
CC	activity, and can be used in	
CC	proteins and polynucleotides	
CC	treatment of diseases associ	
CC	example, they may be used to	
CC	expression by rectifying mut	
CC	that affect the activity of	
CC	supplement the patients own	
CC	polynucleotides may be used	
CC	the nucleic acids into a hos	
CC	protein. (I) proteins and po	
CC	diagnose and treat immune/h	
CC	cancers and cancer metastase	
CC	to AAK87694 represent human	
CC	sequences from the present i	
CC	represent sequences used in	
XX		
SO	Sequence 16605 BP; 4009 A; 3	

Query Match 2.3%; Score 34; DB 22; Length 16605;
Best Local Similarity 52.9%; Pred. No. 34;
Matches 73; Conservative 0; Mismatches 65; Indels 0; Gaps 0;

528 GCTATTTCGATGATAAAAAAACTCTTCCTGTTCTGATTTCTCGCCGCTATCGGGGAGC 587

Db 14753 GATATAAATACAAATAAATTCACCTCTTTCTATATACCATTTCCATGCTTTATGACAAGTG 14812
 Qy 588 TTTTCTACCGCTGTAGCGCTGATAAATAAGAGAGACTCAAAATTTCTACTATCCAGAAACA 647
 Db 14813 TGTCTAAGTTCTATATACTACTACCACAGTTGAGACTTAAATTTCTACTATCTCAAAAAG 14872
 Qy 648 CTGGATTTAACTCCTCTG 665
 Db 14873 TTTCTTAGCCACTTCAG 14890

Search completed: June 18, 2003, 02:09:14
 Job time : 270.397 secs

THIS PAGE BLANK (USPTO)

83	32.4	2.7	11649	22	AAS28910	Human immunoglobul
c 84	32.4	2.7	11650	22	AAK89078	Human digestive sy
c 85	32.4	2.7	11650	22	AAK89077	Human digestive sy
c 86	32.4	2.7	12441	20	AAK06750	Salmonella enteric
c 87	32.2	2.7	561	23	ABV47736	Human prostate exp
c 88	32.2	2.7	564	23	ABV48561	Human prostate exp
c 89	32.2	2.7	1254	22	AAH65416	C glutamicum codin
c 90	32.2	2.7	1272	24	ABQ90218	M. capsulatus gene
c 91	32.2	2.7	1689	24	ABK63827	Rat sequence diffe
c 92	32.2	2.7	3519	16	AAO94352	BHVI gi glycoprote
c 93	32.2	2.7	4383	23	ABLO7815	Drosophila melanog
c 94	32.2	2.7	7178	24	ABR84453	Human cDNA differe
c 95	32.2	2.7	7178	24	ABL61941	Human cDNA differe
c 96	32.2	2.7	7367	24	ABR84452	Human cDNA differe
c 97	32.2	2.7	9721	23	ABLO7586	Drosophila melanog
c 98	32.2	2.7	349980	22	AAH68525	C glutamicum codin
c 99	32	2.7	435	23	ABV49147	Human prostate exp
100	32	2.7	487	23	ABV32296	Human prostate exp
101	32	2.7	487	23	ABV41127	Human prostate exp
102	32	2.7	487	23	ABV44673	Human prostate exp
103	32	2.7	493	23	ABV48448	Human prostate exp
104	32	2.7	601	23	ABV48230	Human prostate exp
c 105	32	2.7	1311	23	AA552027	Staphylococcus aur
c 106	32	2.7	1563	23	AA554470	Staphylococcus aur
c 107	32	2.7	4590	22	AAH24065	Yeast ADP3604-asso
108	32	2.7	50000	21	AAK96367	Polymorphic repeat
109	31.8	2.7	400	23	ABV16382	Human prostate exp
110	31.8	2.7	456	23	ABV47610	Human prostate exp
111	31.8	2.7	479	23	ABV49156	Human prostate exp
112	31.8	2.7	485	23	ABV48591	Human prostate exp
113	31.8	2.7	487	23	ABV48156	Human prostate exp
114	31.8	2.7	502	23	ABV48780	Human prostate exp
115	31.8	2.7	505	23	ABV48972	Human prostate exp
116	31.8	2.7	509	23	ABV49208	Human prostate exp
117	31.8	2.7	509	23	ABV49696	Human prostate exp
118	31.8	2.7	575	23	ABV48521	Human prostate exp
119	31.8	2.7	597	23	ABV46464	Human prostate exp
c 120	31.8	2.7	619	23	ABV46740	Human prostate exp
c 121	31.8	2.7	1146	24	AAJ28444	Bacillus subtilis
c 122	31.8	2.7	3153	22	AAE29464	Murine brevidin cd
123	31.8	2.7	7555	23	ABL10108	Drosophila melanog
124	31.6	2.7	276	23	ABV49839	Human prostate exp
125	31.6	2.7	362	23	ABV49714	Human prostate exp
126	31.6	2.7	410	23	ABV44750	Human prostate exp
127	31.6	2.7	411	23	ABV49771	Human prostate exp
128	31.6	2.7	437	23	ABV48648	Human prostate exp
129	31.6	2.7	449	23	ABV47740	Human prostate exp
130	31.6	2.7	458	23	ABV49161	Human prostate exp
131	31.6	2.7	460	23	ABV46085	Human prostate exp
132	31.6	2.7	460	23	ABV48897	Human prostate exp
133	31.6	2.7	462	23	ABV46017	Human prostate exp
134	31.6	2.7	462	23	ABV49236	Human prostate exp
135	31.6	2.7	467	23	ABV48423	Human prostate exp
136	31.6	2.7	469	23	ABV49626	Human prostate exp
137	31.6	2.7	471	23	ABV47053	Human prostate exp
138	31.6	2.7	474	23	ABV47863	Human prostate exp
139	31.6	2.7	474	23	ABV49335	Human prostate exp
140	31.6	2.7	475	23	ABV48575	Human prostate exp
141	31.6	2.7	476	23	ABV48002	Human prostate exp
142	31.6	2.7	479	23	ABV48777	Human prostate exp
143	31.6	2.7	479	23	ABV49527	Human prostate exp
144	31.6	2.7	481	23	ABV47790	Human prostate exp
145	31.6	2.7	481	23	ABV48948	Human prostate exp
146	31.6	2.7	485	23	ABV46908	Human prostate exp
147	31.6	2.7	490	23	ABV47641	Human prostate exp
148	31.6	2.7	490	23	ABV47702	Human prostate exp
149	31.6	2.7	491	23	ABV49046	Human prostate exp
150	31.6	2.7	492	23	ABV49119	Human prostate exp

ALIGNMENTS

RESULT 1

AAZ36102

ID AAZ36102 standard; DNA; 1181 BP.

XX

AC AAZ36102;

XX

DT 11-FEB-2000 (first entry)

XX

DE Nucleic acid sequence specific to enterohemorrhagic Escherichia coli.

XX

KW Enterohemorrhagic Escherichia coli; EHEC; virulence factor;

KW enterohemolysine; ehly; intimin; eae; virk gene; E. coli O157:H7; ds.

XX

OS Escherichia coli.

XX

PN WO9955908-A2.

XX

PD 04-NOV-1999.

XX

PF 27-APR-1999; 99WO-FR01000.

XX

PR 28-APR-1998; 98PR-0005329.

XX

PA (SNFI) PASTEUR SANOPI DIAGNOSTICS.

XX

PI Frechon DTM, Laure FC, Thierry D;

XX

PS WPI; 2000-013443/01.

XX

CC New nucleic acid containing sequences specific to enterohemorrhagic

CC Escherichia coli, particularly serotype O157:H7, used for detecting

CC these bacteria in food

CC

CC Claim 1; Fig 2; 48pp; French.

CC

CC The present sequence is specific to enterohemorrhagic Escherichia coli

CC (EHEC). The sequence associated with the presence of virulence

CC factors enterohemolysine (ehly) and intimin (eae). Nucleotides 237-570

CC also have 88% homology with the virk gene which codes for virulence

CC proteins of Shigella flexneri. The present sequence is of plasmid origin.

CC Fragments of the present sequence are used, as probes and primers, for

CC detection of E. coli O157:H7 and other enterohemorrhagic E. coli (EHEC),

CC in human or animal samples, foods or the environment. The fragments are

CC also useful for epidemiological studies.

CC

XX

SQ Sequence 1181 BP; 305 A; 317 C; 277 G; 282 T; 0 other;

Query Match 100.0%; Score 1181; DB 21; Length 1181;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1181; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGCAGGAGATGGAAGGCAAAATGAAATGAAATGAAATGAAATGAAATGAAATG 60

Db 1 CTGCAGGAGATGGAAGGCAAAATGAAATGAAATGAAATGAAATGAAATG 60

QY 61 AAGTAGGCGCTGTTCTGTCGGTATTATTAATGCATGACCGTCCCGCTTTAAACAATG 120

Db 61 AAGTAGGCGCTGTTCTGTCGGTATTATTAATGCATGACCGTCCCGCTTTAAACAATG 120

QY 121 TGAATAATTACTCCGTTACCGGAAACCCCTGAAACAAATTCGGGTGAAAGAGATCC 180

Db 121 TGAATAATTACTCCGTTACCGGAAACCCCTGAAACAAATTCGGGTGAAAGAGATCC 180

QY 181 GCGGTTATCTGTTGCTATTTCCCTTAGCTGACTGACCCAGACACAATGATCTGTGCGG 240

Db 181 GCGGTTATCTGTTGCTATTTCCCTTAGCTGACTGACCCAGACACAATGATCTGTGCGG 240

QY 241 TTCTGTATATCAAAACCGGTACTCAATATCTTCTGCGGTGGTGGCCATCCCGGA 300

Db 241 TTCTGTATATCAAAACCGGTACTCAATATCTTCTGCGGTGGTGGCCATCCCGGA 300

QY 301 AGCGTTCCGTCGGGATGAAATGCGCAGTCCGCGGTCCATGACAGACATCCCCCA 360

301	AGCGTTCGGGTCCGGGATAAAAAATCGCGCAGTCGCGCGGTTCATGACAGACACATATCCCCA	360
361	CGGGTAACAGCGTCCCTGTGCACATTTCTTGAATGACATCAGGGATCCCCCGCGTCTCAC	420
361	CGGGTAACAGCGTCCCTGTGCACATTTCTTGAATGACATCAGGGATCCCCCGCGTCTCAC	420
421	TGGCGATAAGCGGCACGCGCGGAGACTGACGCTTCAGCCAGTACCATACCAACGCTTCAT	480
421	TGGCGATAAGCGGCACGCGCGGAGACTGACGCTTCAGCCAGTACCATACCAACGCTTCAT	480
481	TTTCCGAAGGCATGACCACACACTGSCAATCCGGTAGCCGGTAAACGCTGGGAAAAGGG	540
481	TTTCCGAAGGCATGACCACACACTGSCAATCCGGTAGCCGGTAAACGCTGGGAAAAGGG	540
541	CACCTGCCATTACACATCTCCGGTCAATCCCAAGGTGTTCTGCTGTCGACGACGACGTG	600
541	CACCTGCCATTACACATCTCCGGTCAATCCCAAGGTGTTCTGCTGTCGACGACGACGTG	600
601	CTTTCGTAATCTTCACGCCCGCGGCCACACAGCAGCAGCGAATGATTTCCCTTCCATCT	660
601	CTTTCGTAATCTTCACGCCCGCGGCCACACAGCAGCAGCGAATGATTTCCCTTCCATCT	660
661	TCAGCTGATACAAATACACGACGATAAATTCATGTCCTTTTTCGGGACGTAGCATCCCCA	720
661	TCAGCTGATACAAATACACGACGATAAATTCATGTCCTTTTTCGGGACGTAGCATCCCCA	720
721	CCTGAACGATAAGCGGNACATTTCTGCTGATCGACGCCAGGCGTGGAATGACAGGGGTA	780
721	CCTGAACGATAAGCGGAACATTTCTGCTGATCGACGCCAGGCGTGGAATGACAGGGGTA	780
781	ACGGTCGATGGCTTCATTATGCAATCGCGGCCAGTCGAAACCCGGTGGGAATAACCGTTA	840
781	ACGGTCGATGGCTTCATTATGCAATCGCGGCCAGTCGAAACCCGGTGGGAATAACCGTTA	840
841	CCGGTGTCTGACACCTTCGCCCATCAGATGCGCCCATCATGGTGAGATAGGCACACAA	900
841	CCGGTGTCTGACACCTTCGCCCATCAGATGCGCCCATCATGGTGAGATAGGCACACAA	900
901	TGAAATCACACAGATAATTCAGGGAAAACGTTCTGGTCTTACGGGTGATCTAGGTTTTTT	960
901	TGAAATCACACAGATAATTCAGGGAAAACGTTCTGGTCTTACGGGTGATCTAGGTTTTTT	960
961	GTCTGACAAATAGTGAAGCGGTGACAGCATATCAGACGGCTCAGTCTGCTATATTACTGT	1020
961	GTCTGACAAATAGTGAAGCGGTGACAGCATATCAGACGGCTCAGTCTGCTATATTACTGT	1020
1021	CATGGCCACTATGCGAGATCACCAGATCAGGTTTAAATTCGCCGATAATCCGTCGAGTC	1080
1021	CATGGCCACTATGCGAGATCACCAGATCAGGTTTAAATTCGCCGATAATCCGTCGAGTC	1080
1081	TGAGGATGGAAGGAAGGTGAAGCGTGTCTCCTGAAAGGAATAAAAGTGACATCATGCCCTC	1140
1081	TGAGGATGGAAGGAAGGTGAAGCGTGTCTCCTGAAAGGAATAAAAGTGACATCATGCCCTC	1140
1141	TTTTTCTCGGCTCCGGAGCAATTTTACTTTTTTCTCTGCAAG	1181
1141	TTTTTCTCGGCTCCGGAGCAATTTTACTTTTTTCTCTGCAAG	1181

RESULT 2
AAQ55982/C
ID AAQ55982 standard; DNA; 1642 BP.

XX DE Plasmid pAN240 carrying the B. subtilis PanBCD operon.
 XX DE PanBCD operon: pantothenate biosynthesis; cyclic; circular; pAN240;
 KW vitamin B5; nutritional supplement; panto-compound; pantoate; ds.
 XX OS Chimeric - Bacteriophage SP01.
 OS Chimeric - Escherichia coli.
 OS Chimeric - Bacillus subtilis.
 XX PN WO200121772-A2.
 XX PD 29-MAR-2001.
 XX PF 21-SEP-2000; 2000WO-US25993.
 XX PR 21-SEP-1999; 99US-0400494.
 PR 07-JUN-2000; 2000US-0210072.
 PR 28-JUL-2000; 2000US-0221836.
 PR 24-AUG-2000; 2000US-0227860.
 XX PA (OMNI-) OMNIGENE BIOPRODUCTS.
 XX PI Yocum RR, Patterson TA, Hermann T, Pero JG;
 XX WPI; 2001-218644/22.
 XX PT New recombinant microorganism which overexpress a Bacillus subtilis
 PT pantothenate biosynthetic enzyme, useful for the high yield production
 PT of panto-compounds such as pantothenate and pantoate -
 XX Example 1; Page 232-238; 292pp; English.
 XX The sequence, pAN240, is a plasmid carrying the B. subtilis PanBCD operon
 CC which encodes three enzymes of the pantothenate biosynthetic pathway
 CC (PanB, C and D). Pantothenate, also known as vitamin B5, is used as a
 CC nutritional supplement in mammals and humans. The invention concerns
 CC methods of producing recombinant microorganisms overexpressing at least
 CC one B. subtilis pantothenate biosynthetic enzyme. The microorganisms and
 CC methods of producing them are useful for producing a panto-compound such
 CC as pantothenate or pantoate, which is a nutritional requirement for
 CC livestock and humans. The methods are also useful for the identification
 CC of pantothenate kinase modulators. Panto-compounds are produced at a
 CC significantly higher yield than prior art methods and can be produced
 CC independent of the need to feed precursors which decreases expense.
 XX Sequence 10801 BP; 2797 A; 2486 C; 2151 G; 3367 T; 0 other;
 SQ
 Query Match 5.4%; Score 64.2; DB 22; Length 10801;
 Best Local Similarity 54.4%; Pred. No. 1.2e-09;
 Matches 129; Conservative 0; Mismatches 108; Indels 0; Gaps 0;
 QY 335 GCGGTCCATCAGACATCCGCCAGGGTAACAGCGTCCCTGTACATTTCTTGAAT 394
 DB 6490 GCCTGTCGGCGGTAACTACCGACATCCACCACCAAAATCCGCTCACATGTTTTAT 6549
 QY 395 GACATCAGGATCCGCGGCTTCACCTGGCGGTAACCGGACGCGCGGAGACTGACGCTTC 454
 DB 6550 AACCTCAGGATACGGCAATGTTTGTTCATTAACAGGCACTCCGCAAGCCATCGCTTC 6609
 QY 455 AGCCAGTACCATACCAACGCTTCATTTCCGAGGAGTACACCACTGGCAATCCG 514
 DB 6610 AAGCAGGACAGGCAAGCTTTCTTTTCAGATAGCAGGCTTCAATTCGCTTAATAGA 6669
 QY 515 GTAGACCGGTAAACGCTGGGAAAGGCGGCACTGCCATTAAACATCTCCGCTCATTTCC 571
 DB 6670 ATAAAGATCTTCAACACGCTTTCTTATTTCCAGCATTAAGACTTGGTCTTCCAGGCC 6726
 RESULT 4
 AAI99683
 ID AAI99683 standard; DNA; 4403765 BP.
 XX

AC AAI99683;
 XX 15-JAN-2002 (first entry)
 XX Mycobacterium tuberculosis strain H37Rv genome SEQ ID NO 2.
 DE Mycobacterium tuberculosis; strain H37Rv; strain CDC 1551; genome;
 KW variation; epidemiology; patient treatment; epidemic monitoring; ds.
 XX Mycobacterium tuberculosis.
 XX US6294328-B1.
 XX PD 25-SEP-2001.
 XX PF 24-JUN-1998; 98US-0103840.
 XX PR 24-JUN-1998; 98US-0103840.
 XX PA (GENO-) INST GENOMIC RES.
 XX FLeischmann RD, White OR, Fraser CM, Venter JC;
 XX WPI; 2001-647261/74.
 XX Claim 4; SEQ ID NO 2; 3pp + Sequence Listing; English.
 XX The invention relates to evaluating strain variation within and between
 CC different populations of the tuberculosis bacterial pathogen,
 CC Mycobacterium tuberculosis or related Mycobacterium by determining the
 CC nucleotide sequence of the first strain at positions in the complete
 CC sequence of the genome that correspond to positions that differ in the
 CC nucleotide sequences of M. tuberculosis strains CDC 1551 (AAI99683) and
 CC H37Rv (AAI99682). The method is useful for evaluating strain variation of
 CC M. tuberculosis and has valuable application in the fields of
 CC tuberculosis genetics, epidemiology, patient treatment and epidemic
 CC monitoring.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from USPTO
 CC at seqdata.uspto.gov/sequence.html?DocID=6294328B1.
 XX SQ Sequence 4403765 BP; 757105 A; 1447799 C; 1441301 G; 757371 T; 189 other;
 Query Match 4.8%; Score 57.2; DB 22; Length 4403765;
 Best Local Similarity 54.2%; Pred. No. 8.3e-06;
 Matches 116; Conservative 0; Mismatches 98; Indels 0; Gaps 0;
 QY 272 TTCTCTGGCGCTGCTGCCATCATCCGGAAGGCTTCGGTCCGGATATAAAATTCGCGAG 331
 DB 2447373 TTCTCGCGCGGCGGCCCATCCGACGCGCCGATCCGATCAGTACGACACTCGGCCAC 2447432
 QY 332 TGGCGCGGTCATGACAGACATCCGCCAGGGTAACAGCGTCCCTGTCTCATTTCTTG 391
 DB 2447433 GGGCTCGGCCACCGCGGTCCACCGACCTACCGTCCGACACTAGCCCAAGTCTTGTGCTG 2447492
 QY 392 AATGACATCAGGATCCCGCCGCTCTCACTGGCGATAACGGCACCGCGGAGACTGACGC 451
 DB 2447493 CACCGTTCCGGCGCTCCGCCAGANTGCCGGGATATCCGGCACCGCGCGGAGGC 2447552
 QY 452 TTCAGCAGTACCATACCAACAGCTTCAATTTTC 485
 DB 2447553 TTCAGGAAACAGCATGCCCAAGCCCTCGACGTCC 2447586
 RESULT 5
 AAI99682
 ID AAI99682 standard; DNA; 4411529 BP.
 XX

```

AC AAI99682;
XX
DT 15-JAN-2002 (first entry)
XX
DE Mycobacterium tuberculosis strain H37Rv genome SEQ ID NO 1.
XX
KW Mycobacterium tuberculosis; strain H37Rv; strain CDC 1551; genome;
XX variation; epidemiology; patient treatment; epidemic monitoring; ds.
XX
OS Mycobacterium tuberculosis.
XX
PN US6294328-B1.
XX
PD 25-SEP-2001.
XX
PF 24-JUN-1998; 98US-0103840.
XX
PR 24-JUN-1998; 98US-0103840.
XX
PA (GENO-) INST GENOMIC RES.
XX
PI Fleischmann RD, White OR, Fraser CM, Venter JC;
XX WPI; 2001-647261/74.
XX
DR Evaluating strain variation of Mycobacterium tuberculosis, comprises
XX determining the nucleotide sequence of the strain at positions in the
XX genome corresponding to positions where M. tuberculosis strains CDC
XX 1551 and H37Rv differ.
XX
PS Claim 3; SEQ ID NO 1; 3pp + Sequence Listing; English.
XX
CC The invention relates to evaluating strain variation within and between
CC different populations of the tuberculosis bacterial pathogen,
CC Mycobacterium tuberculosis or related Mycobacterium by determining the
CC nucleotide sequence of the first strain at positions in the complete
CC sequence of the genome that correspond to positions that differ in the
CC nucleotide sequences of M. tuberculosis strains CDC 1551 (AAI99683) and
CC H37Rv (AAI99682). The method is useful for evaluating strain variation of
CC M. tuberculosis and has valuable application in the fields of
CC tuberculosis genetics, epidemiology, patient treatment and epidemic
CC monitoring.
CC
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from USPTO
CC at seqdata.uspto.gov/sequence.html?DocID=6294328B1.
XX
SQ Sequence 4411529 BP; 758565 A; 1449983 C; 1444602 G; 758379 T; 0 other;
Query Match 4.8%; Score 57.2; DB 22; Length 4411529;
Best Local Similarity 54.2%; Pred. No. 8.3e-06;
Matches 116; Conservative 0; Mismatches 98; Indels 0; Gaps 0;
QY 272 TTCTCTGGCGTGGCTGCCATCATCCGGAGCGTTCCGGTCCGGATATAAAATCGGCGAG 331
DB 2450072 TTCTCGCGCGGGCCCATCCGAGCGCCGATCCGATCGATCAGCACTCGGCCAC 2450131
QY 332 TGGCGCGGTCATGCAGACATATCCCGCCACGGGTAAACGGTCCCTGTGCATATCTCTG 391
DB 2450132 GCGCTCGGCCACCGCGTCCACCGACCTACCGTCGACCTAGCCAGTCTTGTGTGTG 2450191
QY 392 AATGACATCAGGATCCCGCGCTCTCACTGGGATACGGGACCGCGGAGCTACGCG 451
DB 2450192 CACCGTTTCGGCGTCCCGCCAGAAATGCGCGGATATCCGGACCGCGCGGAGGC 2450251
QY 452 TTCAGCGAGTACCATCAAAAGCTTCATTTC 485
DB 2450252 TTCGAGACACAGATCCCGCCAGCCCTCGAGTCC 2450285
RESULT 6
ID ABR74404/c
ID ABR74404 standard; DNA; 1135 BP.
XX

```

```

AC ABK74404;
XX
DT 13-AUG-2002 (first entry)
XX
DE Bacillus licheniformis genomic sequence tag (GST) #1695.
XX
KW Differential gene expression; genomic sequenced tag; GST;
XX altered culture condition; environmental stress;
XX physiological provocation; ds.
XX
OS Bacillus licheniformis.
XX
PN WO200229113-A2.
XX
PD 11-APR-2002.
XX
PF 05-OCT-2001; 2001WO-US31437.
XX
PR 06-OCT-2000; 2000US-0680598.
XX
PR 27-MAR-2001; 2001US-279526P.
XX
PA (NOVO ) NOVOZYMES BIOTECH INC.
XX (NOVO ) NOVOZYMES AS.
XX
PI Berka R, Clausen IG;
XX
DR WPI; 2002-416684/44.
XX
PT Monitoring differential expression of several genes in first Bacillus
PT cell relative to expression of same genes in one or more second
PT Bacillus cells, by using substrate containing Bacillus genomic
PT sequenced tag array.
XX
PS Claim 4; SEQ ID NO 1695; 200pp; English.
XX
CC The invention describes a method of monitoring differential expression of
CC genes in a first Bacillus cell relative to expression of the genes in
CC other Bacillus cells, comprising hybridising labelled nucleic acid probes
CC isolated from Bacillus cells to a substrate containing array of Bacillus
CC genomic sequenced tags (GST), examining the array, and determining
CC relative gene expression by an observed hybridisation reporter signal of
CC a spot in the array. The method is useful for measuring the expression of
CC genes in a first Bacillus cell relative to expression of the same genes
CC in one or more second Bacillus cells. The method is useful for monitoring
CC global expression of several genes from a Bacillus cell, discovering new
CC genes, identifying possible functions of unknown open reading frames and
CC monitoring gene copy number variation and stability. Monitoring changes
CC in expression of genes may be used to provide a representation of the way
CC in which Bacillus cells adapt to changes in culture conditions,
CC environmental stress or other physiological provocation. Extensive
CC follow-up characterisation is unnecessary, when one spot on an array
CC equals one gene or one open reading frame, since sequence information is
CC available. This sequence represents a genomic sequence tag (GST) used in
CC the method of the invention.
CC
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 1135 BP; 327 A; 251 C; 278 G; 279 T; 0 other;
Query Match 4.1%; Score 49; DB 24; Length 1135;
Best Local Similarity 50.2%; Pred. No. 3.5e-05;
Matches 121; Conservative 0; Mismatches 120; Indels 0; Gaps 0;
QY 351 ACATCCCCCAGGGTAACAGCGTCCCTGTGCACATTTCTTGTGATGATCAGGATCCCG 410
DB 977 ATATCTCCGACATCGACCAAGTAACCGGATATCCGTTTGTGATCCTCGGGAATCCCC 918
QY 411 CCGGTCTACTGGCGATAACGGCAGCGGAGACTGACCTTCACCGAGTACCATACCA 470
DB 917 CCGATATTGTGCCGATGACGAGCAGCGCCGATCGCTTCAAGCAGCAGCTAAACCG 858

```

QY 471 AACGCTTCATTTCGGAAGCATGACACACACACTGCGAATCGGTAGACCGGTAAACGCT 530
 DB 857 AAGCTTTCCTTCGGAAGCATGACACACACACTGCGAATCGGTAGACCGGTAAACGCTTCCTACG 798
 QY 531 GGGAAAGGCGACCTGCGCATTAACACATATCCGCTCATTCGCCAGGTGTTCTGTCTGCTGA 590
 DB 797 CTGCTTTGTTGCGCAAGGAAAGACGCTGCTCTTTTAATCCGAGCTGTCTGACAAAGCTGG 738
 QY 591 C 591
 DB 737 C 737

RESULT 7
 ABK78827/c
 ID ABK78827 standard; DNA; 783 BP.
 XX AC
 XX ABK78827;
 XX 13-AUG-2002 (first entry)
 XX Bacillus clausii genomic sequence tag (GST) #1670.
 XX Differential gene expression; genomic sequenced tag; GST;
 KW altered culture condition; environmental stress;
 KW physiological provocation; ds.
 XX Bacillus clausii.
 OS Bacillus clausii.

PN WO200229113-A2.
 PD 11-APR-2002.
 PF 05-OCT-2001; 2001WO-US31437.
 XX 06-OCT-2000; 2000US-0680598.
 PR 27-MAR-2001; 2001US-279526P.
 XX (NOVO) NOVOZYMES BIOTECH INC.
 PA (NOVO) NOVOZYMES AS.
 PI Berka R, Clausen IG;
 XX WPI; 2002-416684/44.
 DR Monitoring differential expression of several genes in first Bacillus
 XX cell relative to expression of same genes in one or more second
 PT Bacillus cells, by using substrate containing Bacillus genomic
 PT sequenced tag array -
 XX Claim 11; SEQ ID NO 6118; 200pp; English.

XX The invention describes a method of monitoring differential expression of
 XX genes in a first Bacillus cell relative to expression of the genes in
 CC other Bacillus cells, comprising hybridising labelled nucleic acid probes
 CC isolated from Bacillus cells to a substrate containing array of Bacillus
 CC genomic sequenced tags (GST), examining the array, and determining
 CC relative gene expression by an observed hybridisation reporter signal of
 CC a spot in the array. The method is useful for measuring the expression of
 CC genes in a first Bacillus cell relative to expression of the same genes
 CC in one or more second Bacillus cells. The method is useful for monitoring
 CC global expression of several genes from a Bacillus cell, discovering new
 CC genes, identifying possible functions of unknown open reading frames and
 CC monitoring gene copy number variation and stability. Monitoring changes
 CC in expression of genes may be used to provide a representation of the way
 CC in which Bacillus cells adapt to changes in culture conditions,
 CC environmental stress or other physiological provocation. Extensive
 CC follow-up characterisation is unnecessary, when one spot on an array
 CC equals one gene or one open reading frame, since sequence information is
 CC available. This sequence represents a genomic sequence tag (GST) used in
 CC the method of the invention.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO

CC at
 CC ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 783 BP; 225 A; 149 C; 199 G; 208 T; 2 other;
 Query Match 4.1%; Score 48.6; DB 24; Length 783;
 Best Local Similarity 49.8%; Pred. No. 3.8e-05;
 Matches 123; Conservative 0; Mismatches 124; Indels 0; Gaps 0;
 QY 354 TCCGCCACGGGTAACAGCGTCCCTGTCACATCTTCTGAATGACATCAGGGATCCCGCC 413
 DB 631 TTCCCAAGGGGCATATAAATACCACTTCTTCAACAACTTCGGGATGCCACCA 572
 QY 414 GTCTCAGTGGCGATAACGGGACGCGGAGACTGACGTTTACGCGAGTACCATACCAAC 473
 DB 571 ATTTTGTGCGCATCACAGGTACCCGCGCATGCGTTCAGGCGCCACTAAACCAAG 512
 QY 474 GCTTCATTTTCCGAGGGCATGACCAACACACATCTCGCTCAATCCGCTAGACCGGTAAACGCTGG 533
 DB 511 CTTCCTTTTTCGCTTAAACAACAGCATTAATCGCTCATAGACAACAATTCAGCAACATGC 452
 QY 534 AAAAGGCGACCTGCCATTAAACACATCTCGCTCAATCCCGAGGTGTTCTGTCTGCTGACGC 593
 DB 451 TTTTGTGCTGCTAGCATAGGACATGCTTTGTCAAGCCGCTTTCTTCCACAAACTGCCTC 392
 QY 594 AGACGTG 600
 DB 391 GCAATTG 385

RESULT 8
 ABK78820/c
 ID ABK78820 standard; DNA; 783 BP.
 XX AC
 XX ABK78820;
 XX 13-AUG-2002 (first entry)
 XX Bacillus clausii genomic sequence tag (GST) #1663.
 XX Differential gene expression; genomic sequenced tag; GST;
 KW altered culture condition; environmental stress;
 KW physiological provocation; ds.
 XX Bacillus clausii.
 OS Bacillus clausii.
 PN WO200229113-A2.
 PD 11-APR-2002.
 PF 05-OCT-2001; 2001WO-US31437.
 XX 06-OCT-2000; 2000US-0680598.
 PR 27-MAR-2001; 2001US-279526P.
 XX (NOVO) NOVOZYMES BIOTECH INC.
 PA (NOVO) NOVOZYMES AS.
 PI Berka R, Clausen IG;
 XX WPI; 2002-416684/44.
 DR Monitoring differential expression of several genes in first Bacillus
 XX cell relative to expression of same genes in one or more second
 PT Bacillus cells, by using substrate containing Bacillus genomic
 PT sequenced tag array -
 XX Claim 11; SEQ ID NO 6111; 200pp; English.

XX The invention describes a method of monitoring differential expression of
 XX genes in a first Bacillus cell relative to expression of the genes in
 CC other Bacillus cells, comprising hybridising labelled nucleic acid probes
 CC isolated from Bacillus cells to a substrate containing array of Bacillus

CC genomic sequenced tags (GST), examining the array, and determining
 CC relative gene expression by an observed hybridisation reporter signal of
 CC a spot in the array. The method is useful for measuring the expression of
 CC genes in a first Bacillus cell relative to expression of the same genes
 CC in one or more second Bacillus cells. The method is useful for monitoring
 CC global expression of several genes from a Bacillus cell, discovering new
 CC genes, identifying possible functions of unknown open reading frames and
 CC monitoring gene copy number variation and stability. Monitoring changes
 CC in expression of genes may be used to provide a representation of the way
 CC in which Bacillus cells adapt to changes in culture conditions,
 CC environmental stress or other physiological provocation. Extensive
 CC follow-up characterisation is unnecessary, when one spot on an array
 CC equals one gene or one open reading frame, since sequence information is
 CC available. This sequence represents a genomic sequence tag (GST) used in
 CC the method of the invention.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at
 CC ftp.wipo.int/pub/published_pct_sequences.
 XX
 XX Sequence 783 BP; 224 A; 150 C; 204 G; 203 T; 2 other;
 Query Match 3.9%; Score 46.2; DB 24; Length 783;
 Best Local Similarity 59.5%; Pred. No. 0.00024;
 Matches 78; Conservative 0; Mismatches 53; Indels 0; Gaps 0;
 QY 354 TCCCCCAGGGGTAAACAGGTCCTGTCACATTTCTGTAATGACATCAGGATCCGCC 413
 Db 633 TTTCCCAAGGGGCGATAAATAAACAGTTACTTCTTCAACAACTTCCGGGATGCCACCA 574
 QY 414 GTCTCACTGGCGATACGGCGGAGACTGACGCTTCACCCAGTACCATACCAAC 473
 Db 573 ATTTTGTGCGGATACAGTACCGCGATGCCATCCCGATGCCATCCAGCCGACATAACCAAG 514
 QY 474 GCTTCATTTC 484
 Db 513 CTTCCTTTTC 503
 RESULT 9
 ABK74480/c
 ID ABK74480 standard; DNA; 1227 BP.
 XX
 AC ABK74480;
 XX
 DT 13-AUG-2002 (first entry)
 XX
 DE Bacillus licheniformis genomic sequence tag (GST) #1771.
 XX
 KW Differential gene expression; genomic sequenced tag; GST;
 KW altered culture condition; environmental stress;
 KW physiological provocation; ds.
 XX
 OS Bacillus licheniformis.
 XX
 PN WO200229113-A2.
 XX
 PD 11-APR-2002.
 XX
 XX 05-OCT-2001; 2001WO-US31437.
 XX
 PR 06-OCT-2000; 2000US-0680598.
 PR 27-MAR-2001; 2001US-279526P.
 XX
 XX (NOVO) NOVOZYMES BIOTECH INC.
 PA (NOVO) NOVOZYMES AS.
 XX
 XX Berka R, Clausen IG;
 PI
 XX WPI; 2002-416684/44.
 DR
 XX Monitoring differential expression of several genes in first Bacillus
 PT cell relative to expression of same genes in one or more second
 PT

PT Bacillus cells, by using substrate containing Bacillus genomic
 PT sequenced tag array
 XX
 PS Claim 4; SEQ ID NO 1771; 200pp; English.
 XX
 CC The invention describes a method of monitoring differential expression of
 CC genes in a first Bacillus cell relative to expression of the genes in
 CC other Bacillus cells, comprising hybridising labelled nucleic acid probes
 CC isolated from Bacillus cells to a substrate containing array of Bacillus
 CC genomic sequenced tags (GST), examining the array, and determining
 CC relative gene expression by an observed hybridisation reporter signal of
 CC a spot in the array. The method is useful for measuring the expression of
 CC genes in a first Bacillus cell relative to expression of the same genes
 CC in one or more second Bacillus cells. The method is useful for monitoring
 CC global expression of several genes from a Bacillus cell, discovering new
 CC genes, identifying possible functions of unknown open reading frames and
 CC monitoring gene copy number variation and stability. Monitoring changes
 CC in expression of genes may be used to provide a representation of the way
 CC in which Bacillus cells adapt to changes in culture conditions,
 CC environmental stress or other physiological provocation. Extensive
 CC follow-up characterisation is unnecessary, when one spot on an array
 CC equals one gene or one open reading frame, since sequence information is
 CC available. This sequence represents a genomic sequence tag (GST) used in
 CC the method of the invention.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at
 CC ftp.wipo.int/pub/published_pct_sequences.
 XX
 XX Sequence 1227 BP; 368 A; 236 C; 300 G; 323 T; 0 other;
 Query Match 3.5%; Score 41.6; DB 24; Length 1227;
 Best Local Similarity 62.5%; Pred. No. 0.01;
 Matches 65; Conservative 0; Mismatches 39; Indels 0; Gaps 0;
 QY 352 CATCCCCCAGGGTAAACAGGTCCTGTCACATTTCTGTAATGACATCAGGATCCGC 411
 Db 1072 CATCTCTTTAGCGGTACCATACCGGTACATTTGTTCAACGATTTGGGAAGTCCCG 411
 QY 412 CGCTCTCACTGGCGATACGGCGGAGACTGACGCTTCA 455
 Db 1012 CTGCATCGCTGACGATAATAGGAAGTCCGCGTAATTGCGCTTCA 969.
 RESULT 10
 ABQ90082/c
 ID ABQ90082 standard; DNA; 1098 BP.
 XX
 AC ABQ90082;
 XX
 DT 01-OCT-2002 (first entry)
 XX
 DE M. capsulatus gene #67 for DNA array.
 XX
 XX Micro array; gene; ds; differential expression; gene expression.
 XX
 OS Methylococcus capsulatus.
 XX
 PN WO200255655-A2.
 XX
 PD 18-JUL-2002.
 XX
 XX 14-JAN-2002; 2002WO-NO00019.
 PF
 PF 12-JAN-2001; 2001NO-0000235.
 PR 12-JAN-2001; 2001NO-0000239.
 XX
 XX (UNIF-) UNIFOB STIFTELSEN UNIV BERGEN.
 PA (TIGR-) TIGR.
 XX
 XX Birkeland NR, Eidhammer I, Jonassen I, Jensen HB, Lien T;
 PI Lillehaug JR, Lossius I, Eisen JA, Fraser CM, Durkin AS;
 PI Salzberg SL;

XX WPI; 2002-557818/59.

XX Novel DNA array useful for determining differential expression of

XX Methylococcus capsulatus genes, comprises polynucleotides or

XX oligonucleotides representative for a selective number of Methylococcus

XX capsulatus genes

XX Claim 19; Page 81-82; 678pp; English.

XX The invention relates to a novel DNA array giving a representation of a

XX number of Methylococcus capsulatus genes. The method of the invention is

XX useful for determination of the differential expression of the genes of

XX M. capsulatus, and for studying gene expression on a genomic scale and in

XX gene expression assays of M. capsulatus genes. The sequences shown in

XX ABQ90016-ABQ91855 represent M. capsulatus genes for use in arrays of the

XX invention.

XX Sequence 1098 BP; 156 A; 390 C; 374 G; 178 T; 0 other;

XX

Query Match 3.5%; Score 40.8; DB 24; Length 1098;

Best Local Similarity 48.7%; Pred. No. 0.018;

Matches 111; Conservative 0; Mismatches 117; Indels 0; Gaps 0;

QY 274 CTCTGGCGCTGGCTGCCATATCCGGAAGCGTTCCTCCGTCGGGATATAAAATCGCGCAGTG 333

DB 1003 CCGGCGCGCTGGCGCCATGTCGCCGCGCAGGTCCCGATCGTCGAGCAGCGCGGATGG 944

QY 334 CGCGGTCCATGCAGACATCCCGCCAGCGGTACAGCGTCCCTGTACATTTCTTCGAA 393

DB 943 CGGCGCGCAGGCGCGCGCATCGCGGTGGGAGCAGTCCATTTCTCGCGTCCGCCA 884

QY 394 TGACATCAGGATCCCGCGCTCTCACTGCGGATAGCGGCGCAGCGGAGACTGACGCTT 453

DB 893 CGGCTTCGGATCCCGCGCGCGGAGGCGCAGATGGGAGCGCGCGCGCGCGCT 824

QY 454 CAGCCAGTACCATCAACCGCTTCATTTCCGAGGCGATGACCA 501

DB 823 GCAGCAGTGAGATGCCAGGCGCTTCGCGCAGGCGCGGATGATACCA 776

RESULT 11

AAZ45259/c

ID AAZ45259 standard; DNA; 14245 BP.

XX

AC AAZ45259;

XX

XX 27-MAR-2000 (first entry)

XX

DE DNA encoding enzymes involved in exopolysaccharide biosynthesis.

XX

XX Exopolysaccharide; EPS; ESP enzyme; EPS1; EPS2; EPS3; EPS4; EPS5; EPS6;

KW EPS7; EPS8; EPS9; EPS10; EPS11; Lactobacillus helveticus strain LH59;

KW activated D-galactose pyranose; saccharide; beta-glycosyltransferase;

KW undecaprenyl-phosphate-glycosyl-1-phosphate-transferase;

KW alpha-glycosyltransferase; EPS polymerase; glycosyltransferase;

KW phosphofuranose; transporter; food; fermented milk product; yoghurt;

KW cheese; flavour stability; organoleptic property; ss.

XX

OS Lactobacillus helveticus.

XX

XX Key Location/Qualifiers

FT CDS 1052..1729

FT

FT /*tag= a

FT /product= "EPS1"

FT /note= "encodes AAY54075"

FT 1733..2848

FT /*tag= b

FT /product= "EPS2"

FT /note= "encodes AAY54076"

FT 2851..3942

FT /*tag= c

FT /product= "EPS3"

FT

FT CDS

FT 3929..5083

FT /*tag= d

FT /product= "EPS4"

FT /note= "encodes AAY54078"

FT 5076..6095

FT /*tag= e

FT /product= "EPS5"

FT /note= "encodes AAY54079"

FT 6098..7090

FT /*tag= f

FT /product= "EPS6"

FT /note= "encodes AAY54080"

FT 7095..8258

FT /*tag= g

FT /product= "EPS7"

FT /note= "encodes AAY54081"

FT 8283..9122

FT /*tag= h

FT /product= "EPS8"

FT /note= "encodes AAY54082"

FT 9135..10253

FT /*tag= i

FT /product= "EPS9"

FT /note= "encodes AAY54083"

FT 10250..11662

FT /*tag= j

FT /product= "EPS10"

FT /note= "encodes AAY54084"

FT 11664..13181

FT /*tag= k

FT /product= "EPS11"

FT /note= "encodes AAY54085; contains 1 stop codon at

FT nucleotides 12867-12869"

XX

XX WO9962316-A2.

XX

XX 09-DEC-1999.

XX

XX 22-APR-1999; 99WO-EP02841.

XX

XX 22-APR-1998; 98EP-0201310.

XX 22-APR-1998; 98EP-0201311.

XX 22-APR-1998; 98EP-0201312.

XX

XX (NEST) SOC PROD NESTLE SA.

XX

XX Stinglele F, Germond JE, Lamothe G;

XX

XX WPI; 2000-097267/08.

XX P-PSDB; AAY54075; AAY54076, AAY54077, AAY54078, AAY54079, AAY54080,

XX AAY54081; AAY54082, AAY54083, AAY54084; AAY54085.

XX

XX New recombinant enzymes for synthesis of exopolysaccharides,

XX particularly in lactic acid bacteria, for improving properties of

XX fermented milk products

XX

XX Claim 11; Page 110-117; 162pp; French.

XX

XX The present sequence encodes enzymes involved in the biosynthesis of

XX exopolysaccharides (EPS). These enzymes are designated EPS1-EPS11, and

XX are encoded by open reading frames eps1-eps11. The enzymes are isolated

XX from Lactobacillus helveticus strain LH59. The proteins are used

XX in a method for the synthesis of EPS, which includes at least one step

XX of forming a bond (alpha or beta-isomer) between C-1 (carrying the

XX reducing aldehyde function, of an activated D-galactose pyranose), and

XX a phosphate on a lipophilic or proteinaceous primer. Stepwise synthesis

XX of EPS occurs with, in each step, addition of a new sugar unit, through

XX its hemi-acetyl function, to an alcoholic hydroxyl of a second sugar

XX unit, present at the end of a chain of sugar residues bonded to the

XX primer. EPS1 has homology with an undecaprenyl-phosphate-glycosyl-1-

XX phosphate-transferase; EPS2, EPS3 and EPS4 are alpha-

XX glycosyltransferases; EPS5 and EPS6 have homology with a beta-

CC	glycosyltransferase; EPS7 has homlogy with an EPS polymerase; EPS8
CC	is a glycosyltransferase; EPS9 catalyses the transfer of phosphoraneose
CC	onto the following repetitive unit; EPS10 transports the polysaccharides
CC	formed; and EPS11 is involved in the synthesis and export of formed
CC	polysaccharides. The EPS enzyme are used to improve properties of foods,
CC	particularly fermented milk products such as yoghurt and cheese,
CC	e.g. their organoleptic properties and flavour stability.
XX	
SQ	Sequence 14245 BP; 5095 A; 1755 C; 2545 G; 4850 T; 0 other;
Query Match 3.4%; Score 40.6; DB 21; Length 14245;	
Best Local Similarity 46.3%; Pred. No. 0.093;	
Matches 133; Conservative 0; Mismatches 154; Indels 0; Gaps 0;	
QY	226 CAATGATCTCTGCGGTTCGTTAATCAACACGGTACTCAATATCTTCTCTGGCGGTG 285
DB	2822 CACTATTTTGTGGGACTCTTTTACCATCAACTTATTTTTTACAAATATTACTTTTCAGTTT 2763
QY	286 CTGCCATCATCCGGAGCGTTCGGTGGGATAAAAATCGCAGCTGCCCGGTCCATG 345
DB	2762 TAACATAAGAGATCTCAAAATCAGGTGTTTAAATAGTTTATCGACTGCATCTGCTGTT 2703
QY	346 CAGACATATCCCCACGGGTAAACAGGTCCTGTCTCACATTTCTTCTGAATCACATCAGGGA 405
DB	2702 TATGGTAATCATCTTTTATCAACTAGCATGCCATTAAACACCATTCGTAACGATATTAGGAA 2643
QY	406 TCCCGGCCGCTCTCACTGGCGGATAACGGGACGCGGGAGCTGACGGTTTCAGCCAGTACCA 465
DB	2642 TAGCATCCACATTTAGTAGCTACATAGGCACTCCACAGCCATATACTCTGGTAGACTA 2583
QY	466 TACCAACGGTTCATTTTCCGAAGGCATGACCCACACACTGGCAATC 512
DB	2582 AACCAAGCCTTCCCAACGAGATAGTACGATCCCATCCCATCAATC 2536

RESULT 12	
AAZ30356/c	
ID	AAZ30356 standard; DNA; 14245 BP.
XX	
AC	AAZ30356;
XX	
DT	11-FEB-2000 (first entry)
XX	
DE	Nucleotide sequence of the eps operon of L. helveticus LH59.
XX	
KW	eps operon; Lactobacillus helveticus LH59; enzyme; eps1; eps2;
KW	eps3; eps4; eps5; eps7; eps8; eps9; eps10; eps11;
KW	xopolysaccharide biosynthesis; EPS; intersugar bond; antitumour;
KW	probiotic; foodstuff; organoleptic quality; flavour;
KW	lactic acid bacteria; acidified milk product; yoghurt; cheese; ss.
XX	
OS	Lactobacillus helveticus.
XX	
Key	Location/Qualifiers
CDS	1052..1729
FT	/*tag= a
FT	/product= "eps1"
FT	/note= "encodes AAY43777; this CDS is specifically
FT	claimed (claim 11)."
CDS	1733..2848
FT	/*tag= b
FT	/product= "eps2"
FT	/note= "encodes AAY43778; this CDS is specifically
FT	claimed (claim 11)."
CDS	2851..3942
FT	/*tag= c
FT	/product= "eps3"
FT	/note= "encodes AAY43779; this CDS is specifically
FT	claimed (claim 11)."
CDS	3929..5083
FT	/*tag= d
FT	/product= "eps4"
FT	/note= "encodes AAY43780; this CDS is specifically
FT	

FT	CDS	5076..6095	claimed (claim 11)";
FT		/*tag= e	
FT		/product= "eps5"	
FT		/note= "encodes AAY43781; this CDS is specifically	
FT		claimed (claim 11)";	
FT	CDS	6098..7090	
FT		/*tag= f	
FT		/product= "eps6"	
FT		/note= "encodes AAY43782; this CDS is specifically	
FT		claimed (claim 11)";	
FT	CDS	7095..8258	
FT		/*tag= g	
FT		/product= "eps7"	
FT		/note= "encodes AAY43783; this CDS is specifically	
FT		claimed (claim 11)";	
FT	CDS	8283..9122	
FT		/*tag= h	
FT		/product= "eps8"	
FT		/note= "encodes AAY43784; this CDS is specifically	
FT		claimed (claim 11)";	
FT	CDS	9135..10253	
FT		/*tag= i	
FT		/product= "eps9"	
FT		/note= "encodes AAY43785; this CDS is specifically	
FT		claimed (claim 11)";	
FT	CDS	10250..11662	
FT		/*tag= j	
FT		/product= "eps10"	
FT		/note= "encodes AAY43786; this CDS is specifically	
FT		claimed (claim 11)";	
FT	CDS	11664..13181	
FT		/*tag= k	
FT		/product= "eps11"	
FT		/note= "encodes AAY43787; the sequence contains 1	
FT		stop codon at nucleotides 12867-12869;	
FT		this CDS is specifically claimed (claim 11)";	
FT	XX		
XX	WO9954475-A2.		
XX			
XX	28-OCT-1999.		
XX			
XX	22-APR-1999;	99WO-EP03011.	
XX			
XX	22-APR-1998;	98EP-0201310.	
PR		98EP-0201311.	
PR	22-APR-1998;	98EP-0201312.	
XX			
PA	(NEST) SOC PROD NESTLE SA.		
XX			
PI	Stingele F, Germond JE, Lamothe G;		
XX			
DR	WPI; 2000-013255/01.		
DR	P-PSDB; AAY43777, AAY43778, AAY43779, AAY43780, AAY43781, AAY43782,		
DR	AAY43783, AAY43784, AAY43785, AAY43786, AAY43787.		
XX			
PT	New recombinant enzymes for biosynthesis of exopolysaccharides having		
PT	e.g. antitumor or probiotic properties or useful in fermented milk		
PT	Products		
XX			
XX	Claim 11; Page 111-118; 163pp; French.		
XX			
CC	The present sequence represents the eps operon of Lactobacillus		
CC	helveticus LH59. The operon contains 11 open reading frames, and		
CC	encodes enzymes (epsi. eps2, eps3, eps4, eps5, eps6, eps7, eps8, eps9		
CC	and eps11) that are involved in the biosynthesis of exopolysaccharides		
CC	(EPS). The enzymes catalyse the formation of specific intersugar bonds		
CC	The enzymes catalyse a process that includes at least one step of		
CC	forming a bond (in alpha or beta anomeric form) between C1, carrying		
CC	the reducing aldehyde group of an activated D-galp (galactose in		
CC	pyranose form), and a phosphate group on a lipophilic or proteinaceous		
CC	primer. The enzymes are used to produce EPS that have antitumor or		
CC	probiotic properties or are used in foodstuffs to improve organoleptic		
CC	properties.		

CC qualities and flavour. When expressed by lactic acid bacteria, EPS impart
CC a free-flowing character and/or a smooth, creamy texture to acidified
CC milk products (yoghurt or cheese).
XX
SQ Sequence 14245 BP; 5095 A; 1755 C; 2545 G; 4850 T; 0 other;
Query Match 3.4%; Score 40.6; DB 21; Length 14245;
Best Local Similarity 46.3%; Pred. No. 0.093;
Matches 133; Conservative 0; Mismatches 154; Indels 0; Gaps 0;
QY 226 CAATGATCTGTGGCTTCTGTTAATATCAAAACGGTACTCAATATCTCTGCGCTGG 285
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2822 CACTATTTTGTGGACACTTATTTTACCACCAAACTATTTTACAAATATTTACTTTCAGTTT 2763
QY 286 CTGCCATCATCCGAAGCGTTCCGGTGGGATAAAATCGCGACGTGCGCGTCCATG 345
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2762 TAACCTAAGAGATCTCAATATAGTGTGTTAATAGTTTATCGATGCAATCTGCTGCTT 2703
QY 346 CAGACACATCCCCCAGCGGTAAACAGCGTCCCTGTCACATTTCTTGAATGACATCAGGGA 405
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2702 TATGTTAATCATCTTTATCACTAGCATGCCATTAACACCATTCGTAACGATATTAGGAA 2643
QY 406 TCCGCGCGTCTACTGGGATACAGGGGAGCGGAGAGCTGAGCTTACGCGAGTACCA 465
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2642 TAGCATCCACATTAGTAGCTACATAGGCACTCCACAAGCCATATATCTGTGTAGAACTA 2583
QY 466 TACCATAAGCGTTTCATTTTCCGAAGGCGATGACACACACACTGGCAATC 512
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2582 AACCAAGCCCTCCACAGGAGATAGTAGCATCCCAACATCCATCAATC 2536

RESULT 13
AAAF25833/C
ID AAF25833 standard; DNA; 130480 BP.
XX
AC AAF25833;
XX
DT 12-APR-2001 (first entry)
XX
DE R. marinus bacteriophage RM387 genomic DNA.
XX
KW Genome; thermophilic enzyme; washing powder; bleaching; ds.
XX
OS Rhodothermus marinus.
XX
PN WO200075335-A2.
XX
PD 14-DEC-2000.
XX
PF 02-JUN-2000; 2000WO-IB00893.
XX
PR 02-JUN-1999; 99US-0137120.
XX
PA (DECO-) DECODE GENETICS EHF.
XX
PI Hjordleifsdottir S, Hreggvidsson GO, Fridjonsson OH, Aevarsson A;
PI Kristjansson JK;
XX
DR WPI; 2001-061727/07.
XX
PT Isolated nucleic acid from bacteriophage RM 378 encoding enzymes useful
PT in recombinant DNA technology
XX
PS Claim 3; Fig 1A-1Q2; 138pp; English.
XX
CC This invention describes a novel isolated nucleic molecule (I) comprising
CC the genome of bacteriophage RM 378. The invention also describes (1) an
CC isolated nucleic acid which encodes a polypeptide obtainable from
CC bacteriophage RM 378, or its active derivative or fragment; (2) an
CC isolated bacteriophage RM 378 (II); (3) a DNA construct (III) comprising
CC operatively linked to a regulatory sequence; (4) a host cell comprising
CC (III); and (5) an isolated polypeptide (IV) obtainable from (II), its
CC active derivative or fragment. Bacteriophage RM 378 is useful for

CC producing thermophilic enzymes useful in DNA research and commercial
CC settings (e.g. proteases and lipases used in washing powder, hydrolytic
CC enzymes used in bleaching). The isolated nucleic acid molecules and
CC vectors are useful in the manufacture of encoded polypeptide, as probes
CC for isolating homologous sequences (e.g. from other bacteriophage in
CC species), as well as for detecting the presence of the bacteriophage in
CC a culture of host cells. The polypeptides can be used as a molecular
CC weight marker on SDS-PAGE gels or on molecular sieve gel filtration
CC columns. Because the host organism of the RM378 bacteriophage is a
CC thermophile, the enzymes and proteins of the RM378 bacteriophage are
CC significantly more thermostable than those of other (e.g. mesophilic)
CC bacteriophages, such as the T4 bacteriophage of Escherichia coli. The
CC enhanced stability of the enzymes and proteins of RM378 bacteriophage
CC allows their use under temperature conditions which would be prohibitive
CC for other enzymes, thus increasing the range of conditions which can be
CC employed not only in DNA research but also in commercial settings.
XX
SQ Sequence 130480 BP; 37304 A; 27752 C; 27687 G; 37737 T; 0 other;
Query Match 3.4%; Score 40.6; DB 22; Length 130480;
Best Local Similarity 60.4%; Pred. No. 0.34;
Matches 67; Conservative 0; Mismatches 44; Indels 0; Gaps 0;
QY 398 ATCAGGGATCCCGCGCTCTCACTGGCGATACGGGACCGCGGAGACTGACGCTCACC 457
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
123773 ATCTGCAATACCAACCGCTCATAGTAGCAACACCGGGGCTCCACACAGCGAGCTCAAG 123714
QY 458 CAGTACCATACCAACGCTTCATTTCCGAAGGCGATGACACACACACTGGC 508
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
123713 CGAAGCGATCCCGGACCCCTCATAGATGAAGCTGCACACACACTCGGC 123663

RESULT 14
AAI99683/C
ID AAI99683 standard; DNA; 4403765 BP.
XX
AC AAI99683;
XX
DT 15-JAN-2002 (first entry)
XX
DE Mycobacterium tuberculosis strain H37Rv genome SEQ ID NO 2.
XX
KW Mycobacterium tuberculosis; strain H37Rv; strain CDC 1551; genome;
KW variation; epidemiology; patient treatment; epidemic monitoring; ds.
XX
OS Mycobacterium tuberculosis.
XX
PN US6294328-B1.
XX
PD 25-SEP-2001.
XX
PF 24-JUN-1998; 98US-0103840.
XX
PR 24-JUN-1998; 98US-0103840.
XX
PA (GENO-) INST GENOMIC RES.
XX
PI Fleischmann RD, White OR, Fraser CM, Venter JC;
XX
DR WPI; 2001-647261/74.
XX
PT Evaluating strain variation of Mycobacterium tuberculosis, comprises
PT determining the nucleotide sequence of the strain at positions in the
PT genome corresponding to positions where M. tuberculosis strains CDC
PT 1551 and H37Rv differ
XX
PS Claim 4; SEQ ID NO 2; 3pp + Sequence Listing; English.
XX
CC The invention relates to evaluating strain variation within and between
CC different populations of the tuberculosis bacterial pathogen, determining the
CC Mycobacterium tuberculosis or related Mycobacterium by determining the
CC nucleotide sequence of the first strain at positions in the complete
CC sequence of the genome that correspond to positions that differ in the

CC nucleotide sequences of M. tuberculosis strains CDC 1551 (AAI99683) and
CC H37Rv (AAI99682). The method is useful for evaluating strain variation of
CC M. tuberculosis and has valuable application in the fields of
CC tuberculosis genetics, epidemiology, patient treatment and epidemic
CC monitoring.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from USPTO
CC at seqdata.uspto.gov/sequence.html?DocID=6294328B1.
XX
SQ Sequence 4403765 BP; 757105 A; 1447799 C; 1441301 G; 757371 T; 189 other;
Query Match 3.4%; Score 39.8; DB 22; Length 4403765;
Best Local Similarity 50.8%; Pred. No. 4.8;
Matches 95; Conservative 0; Mismatches 92; Indels 0; Gaps 0;
QY 333 GCAGCGGTCATGACGACATCCCGCCGTCCTCACTGGCGATACGGGACGGGAGACTGACGCT 392
DB 577989 GCGTCGGCCCACTGACCGACCTCGTCCCGGACACAGGCGTGGCGGTATCCGTCGCGC 577930
QY 393 ATGACATCAGGATCCCGCCGTCCTCACTGGCGATACGGGACGGGAGACTGACGCT 452
DB 577929 ACCGCGAGCGGACGGCCCGCCCGCCCGCGGACCGACCGGTCGCGGACGCTGGGCGC 577870
QY 453 TCAGCCAGTAGTACCATACCAACGCTTCATTTTCGAGAGGATGACCAACACACTGGCAATC 512
DB 577869 TCCACAGCAACAGGCGGAACGACTCGGAGTAGCTGGCGACCGAACCAGGTCGCGCGCC 577810
QY 513 CGGTAGA 519
DB 577809 CGAAACA 577803

RESULT 15
AAI99682/c
ID AAI99682 standard; DNA: 4411529 BP.
XX
AC AAI99682;
XX
DT 15-JAN-2002 (first entry)
XX
DE Mycobacterium tuberculosis strain H37Rv genome SEQ ID NO 1.
XX
KW Mycobacterium tuberculosis; strain H37Rv; strain CDC 1551; genome;
KW variation; epidemiology; patient treatment; epidemic monitoring; ds.
XX
OS Mycobacterium tuberculosis.
XX
PN US6294328-B1.
XX
PD 25-SEP-2001.
XX
XX 24-JUN-1998; 98US-0103840.
XX
XX 24-JUN-1998; 98US-0103840.
XX
PA (GENO-) INST GENOMIC RES.
XX
XX Fleischmann RD, White OR, Fraser CM, Venter JC;
XX WPI; 2001-647261/74.
XX
XX Evaluating strain variation of Mycobacterium tuberculosis, comprises
PT determining the nucleotide sequence of the strain at positions in the
PT genome corresponding to positions where M. tuberculosis strains CDC
PT 1551 and H37Rv differ.
XX
XX Claim 3; SEQ ID NO 1; 3pp + Sequence Listing; English.
XX

XX The invention relates to evaluating strain variation within and between
CC different populations of the tuberculosis bacterial pathogen,
CC Mycobacterium tuberculosis or related Mycobacterium by determining the
CC nucleotide sequence of the first strain at positions in the complete
CC sequence of the genome that correspond to positions that differ in the

CC nucleotide sequences of M. tuberculosis strains CDC 1551 (AAI99683) and
CC H37Rv (AAI99682). The method is useful for evaluating strain variation of
CC M. tuberculosis and has valuable application in the fields of
CC tuberculosis genetics, epidemiology, patient treatment and epidemic
CC monitoring.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from USPTO
CC at seqdata.uspto.gov/sequence.html?DocID=6294328B1.
XX
SQ Sequence 4411529 BP; 758565 A; 1449983 C; 1444602 G; 758379 T; 0 other;
Query Match 3.4%; Score 39.8; DB 22; Length 4411529;
Best Local Similarity 50.8%; Pred. No. 4.8;
Matches 95; Conservative 0; Mismatches 92; Indels 0; Gaps 0;
QY 333 GCAGCGGTCATGACGACATCCCGCCGTCCTCACTGGCGATACGGGAGACTGACGCT 392
DB 576547 GCGTCGGCCCACTGACCGACCTCGTCCCGGACACAGGCGTGGCGGTATCCGTCGCGC 576488
QY 393 ATGACATCAGGATCCCGCCGTCCTCACTGGCGATACGGGAGACTGACGCT 452
DB 576487 ACCGCGAGCGGACGGCCCGCCCGCCCGGACCGGTCGCGGACGCTGGGCGC 576428
QY 453 TCAGCCAGTAGTACCATACCAACGCTTCATTTTCGAGAGGATGACCAACACACTGGCAATC 512
DB 576427 TCCACAGCAACAGGCGGAACGACTCGGAGTAGCTGGCGACCGAACCAGGTCGCGCGCC 576368
QY 513 CGGTAGA 519
DB 576367 CGAAACA 576361

RESULT 16
AAH41227
ID AAH41227 standard; DNA: 265118 BP.
XX
AC AAH41227;
XX
DT 29-OCT-2001 (first entry)
XX
DE Pyrococcus abyssi genomic fragment #6.
XX
KW Hyperthermophilic archaeon; hyperthermophilic protein; ds.
XX
OS Pyrococcus abyssi.
XX
XX Key Location/Qualifiers
FT misc_feature 1..49980
FT /tag= a
FT /note= "This sequence overlaps with the 3' end of
FT AAH41226"
XX
XX FR2792651-A1.
XX
XX 27-OCT-2000.
XX
XX 21-APR-1999; 99FR-0005034.
XX
XX 21-APR-1999; 99FR-0005034.
XX
XX (CNRS) CNRS CENT NAT RECH SCI.
XX (IFRE-) IFREMER INST FR RECH EXPL MER.
XX
XX Forterre P, Thierry JC, Prieur D, Dietrich J, Lecompte O;
XX Querellou J, Weissenbach J, Saurin W, Heilig R;
XX WPI; 2001-126236/14.
XX
XX New nucleotide sequences isolated from Pyrococcus abyssi encode
PT proteins useful in industry.
XX
XX Claim 1; Page 593-665; 1657pp; French.
XX

CC The present invention relates to the genomic sequence of *Pyrococcus*
CC abyssus and *P. abyssi* proteins (see AAB96053-AAB96842). *P. abyssi* is a
CC hyperthermophilic archaeon, which is isolated from deep-sea hydrothermal
CC vents. The present sequence is a fragment of the genomic sequence of *P.*
CC abyssus. The 5' end of this sequence overlaps with the 3' end of AAB41226.
CC The proteins of the present invention have various potential industrial
CC uses, since the proteins are stable at very high temperatures, some up to
CC 110 degrees centigrade.
CC Note: This patent is in the same patent family as WO200065062, which
CC contains additional sequences as shown in AAB99132-AAB99143,
CC AAB75903-AAB75920 and AAG66436.
XX
SQ Sequence 265118 BP; 75226 A; 61302 C; 54005 G; 74585 T; 0 other;

Query Match 3.4%; Score 39.6; DB 22; Length 265118;
Best Local Similarity 52.4%; Pred. No. 1.1;
Matches 87; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

QY 344 TGCAGACATATCCCGGCGGTATACAGCGTCCCTGTGCACATTTCTTGAATGACATCAGG 403
DB 57111 TTTAGCGCTCATCAATGCTATCAACAAGAACCGGTTCTTTCCATCCCTAATCAACTCAGG 57170
QY 404 GATCCCGCGCTCTCATGTGGGTATACGGCGGCGGACGCTGACGCTTCAGCCAGTAC 463
DB 57171 TATCCCGCGCTCTCTTCCAGGCTGGGACAGAGAGGTTCCGCTCAAGGATTAC 57230
QY 464 CATACCAAGCGCTTCATTTCCGAAGCATGACACCACTGCGCA 509
DB 57231 CATGCCAAGCGCTTCCCTCTAGAGGCTAGAGCTAGGCGCAAGTGCA 57276

RESULT 17
AAX84332
ID AAX84332 standard; DNA; 5059 BP.
XX
AC AAX84332;
XX
DT 08-SEP-1999 (first entry)
XX
DE Stealth virus nucleic acid clone, SEQ ID NO: 24.
DE
KW Stealth virus; detection; diagnosis; infection; ss.
XX
OS Stealth virus.
XX
FH Key Location/Qualifiers
FT misc_difference 3605
FT /tag= a
FT /note= "this nucleotide is represented as a * in the
FT specification, and is included to maintain the
FT base numbering given in the specification"
FT
FT misc_difference 3610
FT /tag= b
FT /note= "this nucleotide is represented as a * in the
FT specification, and is included to maintain the
FT base numbering given in the specification"
FT
FT misc_difference 3615
FT /tag= c
FT /note= "this nucleotide is represented as a * in the
FT specification, and is included to maintain the
FT base numbering given in the specification"
FT
FT misc_difference 3630
FT /tag= d
FT /note= "this nucleotide is represented as a * in the
FT specification, and is included to maintain the
FT base numbering given in the specification"
FT
FT misc_difference 3631
FT /tag= e
FT /note= "this nucleotide is represented as a * in the
FT specification, and is included to maintain the
FT base numbering given in the specification"
FT
FT misc_difference 3638
FT /tag= f
FT /note= "this nucleotide is represented as a * in the
FT specification, and is included to maintain the
FT base numbering given in the specification"

FT /note= "this nucleotide is represented as a * in the
FT specification, and is included to maintain the
FT base numbering given in the specification"
FT
FT misc_difference 3641
FT /tag= g
FT /note= "this nucleotide is represented as a * in the
FT specification, and is included to maintain the
FT base numbering given in the specification"
FT
FT misc_difference 3642
FT /tag= h
FT /note= "this nucleotide is represented as a * in the
FT specification, and is included to maintain the
FT base numbering given in the specification"
FT
FT misc_difference 3657
FT /tag= i
FT /note= "this nucleotide is represented as a * in the
FT specification, and is included to maintain the
FT base numbering given in the specification"
FT
FT misc_difference 3659
FT /tag= j
FT /note= "this nucleotide is represented as a * in the
FT specification, and is included to maintain the
FT base numbering given in the specification"
FT
FT misc_difference 3681
FT /tag= k
FT /note= "this nucleotide is represented as a * in the
FT specification, and is included to maintain the
FT base numbering given in the specification"
FT
FT misc_difference 3698
FT /tag= l
FT /note= "this nucleotide is represented as a * in the
FT specification, and is included to maintain the
FT base numbering given in the specification"
FT
FT misc_difference 3703
FT /tag= m
FT /note= "this nucleotide is represented as a * in the
FT specification, and is included to maintain the
FT base numbering given in the specification"
FT
FT misc_difference 3708
FT /tag= n
FT /note= "this nucleotide is represented as a * in the
FT specification, and is included to maintain the
FT base numbering given in the specification"
FT
FT misc_difference 3713
FT /tag= o
FT /note= "this nucleotide is represented as a * in the
FT specification, and is included to maintain the
FT base numbering given in the specification"
FT
FT misc_difference 3720
FT /tag= p
FT /note= "this nucleotide is represented as a * in the
FT specification, and is included to maintain the
FT base numbering given in the specification"
FT
FT misc_difference 3732
FT /tag= q
FT /note= "this nucleotide is represented as a * in the
FT specification, and is included to maintain the
FT base numbering given in the specification"
FT
FT misc_difference 3733
FT /tag= r
FT /note= "this nucleotide is represented as a * in the
FT specification, and is included to maintain the
FT base numbering given in the specification"
FT
FT misc_difference 3740
FT /tag= s
FT /note= "this nucleotide is represented as a * in the
FT specification, and is included to maintain the
FT base numbering given in the specification"
FT
FT misc_difference 3747
FT /tag= t
FT /note= "this nucleotide is represented as a * in the
FT specification, and is included to maintain the
FT base numbering given in the specification"

```
FT misc_difference 3750 /tag- u "this nucleotide is represented as a * in the
FT /note- specification, and is included to maintain the
FT base numbering given in the specification"
FT
FT misc_difference 3751 /tag- v "this nucleotide is represented as a * in the
FT /note- specification, and is included to maintain the
FT base numbering given in the specification"
FT
FT misc_difference 3766 /tag- w "this nucleotide is represented as a * in the
FT /note- specification, and is included to maintain the
FT base numbering given in the specification"
FT
FT misc_difference 3778 /tag- x "this nucleotide is represented as a * in the
FT /note- specification, and is included to maintain the
FT base numbering given in the specification"
FT
FT misc_difference 3792 /tag- y "this nucleotide is represented as a * in the
FT /note- specification, and is included to maintain the
FT base numbering given in the specification"
FT
FT misc_difference 3799 /tag- z "this nucleotide is represented as a * in the
FT /note- specification, and is included to maintain the
FT base numbering given in the specification"
FT
FT misc_difference 3800 /tag- aa "this nucleotide is represented as a * in the
FT /note- specification, and is included to maintain the
FT base numbering given in the specification"
FT
FT misc_difference 3812 /tag- ab "this nucleotide is represented as a * in the
FT /note- specification, and is included to maintain the
FT base numbering given in the specification"
FT
FT misc_difference 3813 /tag- ac "this nucleotide is represented as a * in the
FT /note- specification, and is included to maintain the
FT base numbering given in the specification"
FT
FT misc_difference 3818 /tag- ad "this nucleotide is represented as a * in the
FT /note- specification, and is included to maintain the
FT base numbering given in the specification"
FT
FT misc_difference 3819 /tag- ae "this nucleotide is represented as a * in the
FT /note- specification, and is included to maintain the
FT base numbering given in the specification"
FT
FT misc_difference 3824 /tag- af "this nucleotide is represented as a * in the
FT /note- specification, and is included to maintain the
FT base numbering given in the specification"
FT
FT misc_difference 3827 /tag- ag "this nucleotide is represented as a * in the
FT /note- specification, and is included to maintain the
FT base numbering given in the specification"
FT
FT misc_difference 3841 /tag- ah "this nucleotide is represented as a * in the
FT /note- specification, and is included to maintain the
FT base numbering given in the specification"
FT
FT misc_difference 3853 /tag- ai "this nucleotide is represented as a * in the
FT /note- specification, and is included to maintain the
FT base numbering given in the specification"
```

```
FT misc_difference 3857 /tag- a "this nucleotide is represented as a * in the
FT /note- specification, and is included to maintain the
FT base numbering given in the specification"
FT
FT misc_difference 3860 /tag- ak "this nucleotide is represented as a * in the
FT /note- specification, and is included to maintain the
FT base numbering given in the specification"
FT
FT misc_difference 3863 /tag- al "this nucleotide is represented as a * in the
FT /note- specification, and is included to maintain the
FT base numbering given in the specification"
FT

Query Match 3.3%; Score 39.2; DB 20; Length 5059;
Best Local Similarity 14.5%; Pred. No. 0.15;
Matches 48; Conservative 138; Mismatches 146; Indels 0; Gaps 0;

QY 809 GGCCAGTCGAACCCGGTGAATACCGTTACCGGTGCTGCTGACACCTTCGCCCATCAG 868
Db 1564 SSSHWNTATBHCHMYRTNSWHCHMATCHDTRMTHSTATHVRSNCSARNDCATBYND 1623
QY 869 ATCCGCCATCATGGTTCAGATAGGCACACAATCAATCAATCAATCAATCAATCAATCA 928
Db 1624 RNNGTMTATCHNGRGNSTHCHMYGNMRWHCHNCTDANDAMNDCSNCDATAARVAABRR 1683
QY 929 CGTTCTGGTCTTACGGGTGATGAGTTTCTGCTGACAATAGTGAAGCGGTGACAGCA 988
Db 1684 HSSCYMTGAVRSRCHMVRACHCMARDTTHCRSNDNGSTATHVRSNCSANAYSSSH 1743
QY 989 TATCAGACGGCTCAGTCTCTATATCTGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCT 1048
Db 1744 WDASGNCANTYHGHRBASTNANDBASTHMGTYTHSTATHVRSNCSNTRHCHMVCARDTHCM 1803
QY 1049 AGTTTAAATTCGCCGATATATCGTCAAGTCTGAGATGAGGAGGAGGAGGAGGAGG 1108
Db 1804 VRSWTHATYCHNGSNCSANADDTNACMARNSCNBMDADWTHHGNBANKNTRSRSCMVNTH 1863
QY 1109 CCTGAAGGAATAAAAGTACATCATGCCCTC 1140
Db 1864 SCASSANYNGRATRHMGYWASSNWNTHSCMVTH 1895

RESULT 18
AAH41226/c
ID AAH41226 standard; DNA: 349980 BP.
AC AAH41226;
XX
XX 29-OCT-2001 (first entry)
DT
DE Pyrococcus abyssi genomic fragment #5.
DE Hyperthermophilic archaeon; hyperthermophilic protein; ds.
XX
XX Pyrococcus abyssi.
OS
XX Key Location/Qualifiers
FH misc_feature 1..49980
FT /tag- a
FT /note- "this sequence overlaps with the 3' end of
FT AAH41225"
FT misc_feature 300001..349980
FT /tag- b
FT /note- "this sequence overlaps with the 5' end of
FT AAH41227"
XX
XX FR2792651-A1.
PN
XX 27-OCT-2000.
PD
XX 21-APR-1999; 99FR-0005034.
PF
```

XX 21-APR-1999; 99FR-0005034.
 XX (CNRS) CNRS CENT NAT RECH SCI.
 XX (IFRE-) IFREMER INST FR RECH EXPL MER.
 XX Forterre P, Thierry JC, Prieur D, Dietrich J, Lecompte O;
 PI Querellou J, Weissenbach J, Saurin W, Heilig R;
 XX WPI; 2001-126236/14.
 XX New nucleotide sequences isolated from *Pyrococcus abyssi* encode
 PT proteins useful in industry -
 XX Claim 1; Page 511-606; 1657pp; French.
 XX The present invention relates to the genomic sequence of *Pyrococcus*
 CC *abyssi* and *P. abyssi* proteins (see AAB96053-AAB96842). *P. abyssi* is a
 CC hyperthermophilic archaeon, which is isolated from deep-sea hydrothermal
 CC vents. The present sequence is a fragment of the genomic sequence of *P.*
 CC *abyssi*. The 5' end of this sequence overlaps with the 3' end of AAB41225
 CC and the 3' end of this sequence overlaps with the 5' end of AAB41227. The
 CC proteins of the present invention have various potential industrial uses,
 CC since the proteins are stable at very high temperatures, some up to 110
 CC degrees centigrade.
 CC Note: This patent is in the same patent family as WO2000065062, which
 CC contains additional sequences as shown in AAB99132-AAB99143.
 CC AAB75903-AAH75920 and AAG66436.
 XX Sequence 349980 BP; 98084 A; 80447 C; 77665 G; 93784 T; 0 other;
 SQ

Query Match 3.3%; Score 39.2; DB 22; Length 349980;
 Best Local Similarity 54.1%; Pred. No. 1.7;
 Matches 80; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

QY 333 GCGCGGTCGATGACAGACATCCCGGATACAGCGTCCCTGTTCACATCTTCTGA 392
 DB 224168 GCTTCGGCCAACTCTAAATGGATCTTTAGTGGAACAAGTATCCCGATCTTTCATCTCG 224109

QY 393 ATGACATCAGGATCCCGCGGTCTCACTGGCGATACGGGACGCGGAGACTGACGCT 452
 DB 224108 ATGACATGCTTTATACCTCCGACATTTGTTCCATAACTGCTTCCACTAGCACCAGCT 224049

QY 453 TCAGCGAGTACCATACCAACGCTTCAT 480
 DB 224048 TCTATCAGAACCATTTCCAAACCCCTCTT 224021

RESULT 19
 ABQ90189/c
 ID ABQ90189 standard; DNA; 1110 BP.
 XX AC ABQ90189;
 XX DT 01-OCT-2002 (first entry)
 XX DE M. capsulatus gene #174 for DNA array.
 XX KW Micro array; gene; ds; differential expression; gene expression.
 XX OS Methylococcus capsulatus.
 XX PN WO200255655-A2.
 XX PD 18-JUL-2002.
 XX PF 14-JAN-2002; 2002WO-NO00019.
 XX PR 12-JAN-2001; 2001NO-0000235.
 XX PR 12-JAN-2001; 2001NO-0000239.
 XX PA (UNIF-) UNIFOB STIFTTELSEN UNIV BERGEN.
 XX (TIGR-) TIGR.

XX Birkeland NK, Eidhammer I, Jonassen I, Jensen HB, Lien T;
 PI Lillehaug JR, Lossius I, Eisen JA, Fraser CM, Durkin AS;
 PI Salzberg SL;
 XX WPI; 2002-557818/59.
 XX Novel DNA array useful for determining differential expression of
 PT Methylococcus capsulatus genes, comprises polynucleotides or
 PT oligonucleotides representative for a selective number of Methylococcus
 PT capsulatus genes -
 XX Claim 19; Page 135-136; 678pp; English.
 XX The invention relates to a novel DNA array giving a representation of a
 CC number of Methylococcus capsulatus genes. The method of the invention is
 CC useful for determination of the differential expression of the genes of
 CC M. capsulatus, and for studying gene expression on a genomic scale and in
 CC gene expression assays of M. capsulatus genes. The sequences shown in
 CC ABO90016-ABO91855 represent M. capsulatus genes for use in arrays of the
 CC invention.
 XX Sequence 1110 BP; 160 A; 344 C; 403 G; 203 T; 0 other;
 SQ

Query Match 3.3%; Score 38.4; DB 24; Length 1110;
 Best Local Similarity 50.5%; Pred. No. 0.11;
 Matches 93; Conservative 0; Mismatches 91; Indels 0; Gaps 0;

QY 325 CGCGCAGTGGCGGTCCTATGCAGACATCCCGGATACAGCGTCCCTGTTCACAT 384
 DB 958 CGACAGGACTCGGCCAGTGTCTGACGTGCGCGCGGCCACGACGACCGCATTC 899

QY 385 TCTTCTGAATGACATCAGGATCCCGCGGTCTCACTGGCGATACGGGACGCGGAGA 444
 DB 898 CCTCGCGGAGAGTTCGGGTATGCGCGGACCGCGTCCGACGACGCGGACCGCG 839

QY 445 CTGACGCTTCAGCCAGTACCATACCAACGCTTCATTTCCGAGGATGACCCACAC 504
 DB 838 CCATCGCTTCCAGATGACAGCGGCGGCGCCCTTCGAGTAGGAGGAGGCGGCGCC 779

QY 505 TGGC 508
 DB 778 GGGC 775

RESULT 20
 ABK74439/c
 ID ABK74439 standard; DNA; 1108 BP.
 XX AC ABK74439;
 XX DT 13-AUG-2002 (first entry)
 XX DE Bacillus licheniformis genomic sequence tag (GST) #1730.
 XX KW Differential gene expression; genomic sequenced tag; GST;
 KW altered culture condition; environmental stress;
 KW physiological provocation; ds.
 XX OS Bacillus licheniformis.
 XX PN WO200229113-A2.
 XX PD 11-APR-2002.
 XX PF 05-OCT-2001; 2001WO-US31437.
 XX PR 06-OCT-2000; 2000US-0680598.
 XX PR 27-MAR-2001; 2001US-279526P.
 XX PA (NOVO) NOVOZYMES BIOTECH INC.
 XX PA (NOVO) NOVOZYMES AS.

DT
XX
DE
XX
XX
KW
KW
KW
XX
OS
XX
FH
FT
FT
FT
FT
XX
PN
XX
PD
XX
PF
XX
PR
PR
PR
XX
PA
PI
XX
XX
DR
DR
PT
PT
PT
XX
PS
XX
SQ

18-JUL-2000 (first entry)
Human cytokine receptor zalphalll degenerate DNA.
Cytokine receptor; zalphalll; human; chromosome 16p11.1;
apoptosis; signal transduction; growth factor; cancer; tumour;
infection; gene therapy; diagnosis; ss.
Homo sapiens.
Key
sig_peptide 1..57
mat_peptide 58..1614
/*tag= a
/*tag= b
WO200017235-A2.
30-MAR-2000.
23-SEP-1999; 99WO-US22149.
23-SEP-1998; 98US-0159254.
09-MAR-1999; 99US-0265117.
06-JUL-1999; 99US-0347930.
(ZYMO) ZYMOGENETICS INC.
Presnell SR, Conklin DC, Novak JE, Hammond AK;
WPI; 2000-292825/25.
P-PSDB; AAY79312.
Novel nucleic acid encoding zalphalll polypeptide, useful for treating
e.g. viral infection or tumors, and for identifying ligands that
stimulate cell proliferation
Claim 2(a); Page 150-151; 190pp; English.
The present sequence is that of a degenerate DNA sequence that
encompasses all DNAs encoding zalphalll (see AAY79312), a novel
human class I cytokine receptor that may be involved in an
apoptotic cellular pathway, or is a cell-cell signalling molecule,
growth factor receptor, or extracellular matrix associated protein
with growth factor hormone activity. The degenerate sequence
serves as a template for optimizing expression of polynucleotides
in various cell types and hosts. The invention provides zalphalll
polypeptides, polynucleotides (see also AAZ94533), agonists,
antagonists and antibodies, and methods for their use in the
treatment and diagnosis of conditions associated with altered
zalphalll expression or activity.
Sequence 1614 BP; 209 A; 190 C; 302 G; 204 T; 709 other;

Query Match 3.1%; Score 36.6; DB 21; Length 1614;
Best Local Similarity 30.7%; Pred. No. 0.55;
Matches 74; Conservative 35; Mismatches 132; Indels 0; Gaps 0;

QY 883 GTGAGATAGGCACACAAATGAATCAGATATTAATTCAGGAAACGTTCTGTCTTAC 942
DB 117 GTGGAAYTNCAYCCNWSNACNNTNACNTGCGARGAYCARTAYGARGARYTNA 176
QY 943 GGGTGTAGTGGTGTGTCGACATAGTGAAGCGGTGACACATATCAGACGGCTCA 1002
DB 177 RGAYGARGCNACNWSNTGYNSNTNAYMNGWSNCGNCAYAYCNCNACNAYGCNACNTA 236
QY 1003 GTCCTGCTATATTACTGTCTGCGCCTATGCGAGATGACAGATCAGGTTTAAATTC 1062
DB 237 YACNTGYCAYATGGAYGNTTYTCATTTATGCGNGAYGAYATHTTWSNGTNAAYATHAC 296
QY 1063 CGATATCCGTCGAGTCTGAGGATGAGGAGAGGTGAGGCTCTTCTGGAAGGAATA 1122
DB 297 NGAYCARWSNGGNAAYTAYWSNCARGARTGYGNGNSNTTYTNTYNGCNGARWSNATHA 356

OS
XX
PN
XX
XX
PD
XX
PF
XX
PR
PR
PR
XX
PA
PI
XX
XX
DR
DR
PT
PT
PT
PT
PT
XX
PS
XX
SQ

18-OCT-2001.
03-APR-2001; 2001WO-US10872.
05-APR-2000; 2000US-194731P.
28-JUL-2000; 2000US-222121P.
(ZYMO) ZYMOGENETICS INC.
Sprecher CA, Novak JE, West JW, Presnell SR, Holly RD, Nelson AJ;
WPI; 2002-025898/03.
Novel soluble receptor polypeptides and polynucleotides used as
cytokine antagonist for stimulating ligand activity-induced
proliferation of hematopoietic cells and for suppressing immune
response in a mammal
Claim 10; Page 179; 243pp; English.
The invention relates to an isolated soluble zalphalll cytokine receptor
polypeptide and their cDNA molecules. Zalpha proteins are useful for
inhibiting or antagonising the ligand activity-induced proliferation of
haematopoietic cells and haematopoietic cell progenitors preferably
lymphoid cells which are natural killer cells or cytotoxic T cells.
Zalpha is useful for treating immune and inflammatory disorders, for
reducing proliferation of neoplastic B or T cells, for suppressing an
immune response in a mammal exposed to an antigen or pathogen. Zalpha is
useful for treating diseases that require immune regulation including
autoimmune diseases such as rheumatoid arthritis, multiple sclerosis,
myasthenia gravis, systemic lupus erythematosus (SLE) and diabetes;
asthma, ulcerative colitis, inflammatory bowel disease, Crohn's disease,
sepsis, viral infection (dengue virus infection) and cancer. The present
sequence is human soluble Zalphalll cytokine receptor cDNA degenerate
sequence.
Sequence 654 BP; 116 A; 69 C; 101 G; 86 T; 282 other;

Query Match 3.1%; Score 36.6; DB 24; Length 654;
Best Local Similarity 30.7%; Pred. No. 0.33;
Matches 74; Conservative 35; Mismatches 132; Indels 0; Gaps 0;

QY 883 GTGAGATAGGCACACAAATGAATCAGATATTAATTCAGGAAACGTTCTGTCTTAC 942
DB 60 GTGGAAYTNCAYCCNWSNACNNTNACNTGCGARGAYCARTAYGARGARYTNA 119
QY 943 GGGTGTAGTGGTGTGTCGACATAGTGAAGCGGTGACACATATCAGACGGCTCA 1002
DB 120 RGAYGARGCNACNWSNTGYNSNTNAYMNGWSNCGNCAYAYCNCNACNAYGCNACNTA 179
QY 1003 GTCCTGCTATATTACTGTCTGCGCCTATGCGAGATGACAGATCAGGTTTAAATTC 1062
DB 180 YACNTGYCAYATGGAYGNTTYTCATTTATGCGNGAYGAYATHTTWSNGTNAAYATHAC 239
QY 1063 CGATATCCGTCGAGTCTGAGGATGAGGAGAGGTGAGGCTCTTCTGGAAGGAATA 1122
DB 240 NGAYCARWSNGGNAAYTAYWSNCARGARTGYGNGNSNTTYTNTYNGCNGARWSNATHA 299
QY 1123 A 1123
DB 300 R 300
RESULT 23
ID AAZ94534
XX AAZ94534 standard; DNA; 1614 BP.
XX AAZ94534;
XX

DT 17-OCT-2001 (first entry)
 XX Probe #8926 used to measure gene expression in human placenta sample.
 DE Probe; microarray; human; placenta; antenatal diagnosis;
 KW genetic disorder; ss.
 XX Homo sapiens.
 OS
 XX WO200157272-A2.
 PN 09-AUG-2001.
 XX 30-JAN-2001; 2001WO-US00663.
 XX 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX (MOLE-) MOLECULAR DYNAMICS INC.
 PA Penn SC, Hanzel DK, Chen W, Rank DR;
 PI WPI; 2001-488897/53.
 DR Human genome-derived single exon nucleic acid probes useful for
 XX analyzing gene expression in human placenta -
 XX Claim 25; SEQ ID No 8926; 654pp; English.
 XX The present invention relates to single exon nucleic acid probes (SENP).
 CC The present sequence is one such probe. The probes are useful for
 CC producing a microarray for predicting, measuring and displaying gene
 CC expression in samples derived from human placenta. The probes are useful
 CC for antenatal diagnosis of human genetic disorders.
 XX Sequence 598 BP; 169 A; 109 C; 137 G; 183 T; 0 other;

Query Match 3.1%; Score 36.2; DB 22; Length 598;
 Best Local Similarity 56.2%; Pred. No. 0.42; Mismatches 0; Gaps 0;
 Matches 68; Conservative 0;
 QY 1053 TTAATCCCGGATTAATCCGTCGAAGTCTGAGGATGAAGGAGGCTGTTCTG 1112
 DB 503 TAAATGTACAGGCCACTTGTACAGAGCTCATGATGACAGTGAATGAACACTCATGTC 444
 QY 1113 AAAGGAATAAAGTGACATCATGCGCTCTTTTCTGCTCCGAGGCAATTTTACTTTT 1172
 DB 443 AACAGATAAAGTGACCTAATACATCTCTGACAGGTTTCTTGAGTTTGGATATTTT 384
 QY 1173 T 1173
 DB 383 T 383

RESULT 27
 AAD22970
 ID AAD22970 standard; cDNA; 741 BP.
 XX
 AC AAD22970;
 XX
 DT 26-FEB-2002 (first entry)
 XX Human soluble zalphall cytokine receptor variant degenerate cDNA.
 DE
 KW Human; Zalphall; cytokine receptor; immunosuppressive; cytostatic;
 KW inflammatory disorder; haemostatic; cell proliferation; immune disorder;
 KW autoimmune disease; rheumatoid arthritis; multiple sclerosis; cancer;
 KW myasthenia gravis; systemic lupus erythematosus; SLE; diabetes; asthma;

KW ulcerative colitis; inflammatory bowel disease; Crohn's disease; sepsis;
 XX viral infection; variant; ss.
 OS Homo sapiens.
 XX WO200177171-A2.
 PN 18-OCT-2001.
 XX 03-APR-2001; 2001WO-US10872.
 XX 05-APR-2000; 2000US-194731P.
 PR 28-JUL-2000; 2000US-222121P.
 XX (ZYMO) ZYMOGENETICS INC.
 XX Sprecher CA, Novak JE, West JW, Presnell SR, Holly RD, Nelson AJ;
 PI WPI; 2002-025898/03.
 DR Novel soluble receptor polypeptides and polynucleotides used as
 XX cytokine antagonist for stimulating ligand activity-induced
 XX proliferation of hematopoietic cells and for suppressing immune
 XX response in a mammal
 XX Disclosure; Page 222-223; 243pp; English.
 PS The invention relates to an isolated soluble zalphall cytokine receptor
 XX polypeptide and their cDNA molecules. Zalphall proteins are useful for
 CC inhibiting or antagonising the ligand activity-induced proliferation of
 CC haematopoietic cells and haematopoietic cell progenitors preferably
 CC lymphoid cells which are natural killer cells or cytotoxic T cells.
 CC Zalphall is useful for treating immune and inflammatory disorders, for
 CC reducing proliferation of neoplastic B or T cells, for suppressing an
 CC immune response in a mammal exposed to an antigen or pathogen. Zalphall is
 CC useful for treating diseases that require immune regulation including
 CC autoimmune diseases such as rheumatoid arthritis, multiple sclerosis,
 CC myasthenia gravis, systemic lupus erythematosus (SLE) and diabetes;
 CC asthma, ulcerative colitis, inflammatory bowel disease, Crohn's disease,
 CC sepsis, viral infection (dengue virus infection) and cancer. The present
 CC sequence is human soluble zalphall cytokine receptor variant cDNA
 CC degenerate sequence.
 XX Sequence 741 BP; 133 A; 79 C; 116 G; 91 T; 322 other;

Query Match 3.1%; Score 36.2; DB 24; Length 741;
 Best Local Similarity 27.6%; Pred. No. 0.48;
 Matches 83; Conservative 48; Mismatches 170; Indels 0; Gaps 0;
 QY 833 AACCGTTACGGGTGCTGACACCTTCGGCCATCAGATCGCCATCATGGTGAGATAGG 892
 DB 139 AAYTNCAYCCNWSNACNTNACNTNACNTGATHTNWSNAAAYACNGGNTGYAT 198
 QY 893 CACAACAATGAATACACAGATAATTAACAGGAAAAACGTTCTGCTTACGGGTGATGA 952
 DB 199 ATHAARGAYMGACNTNGAYTNGCARGAYCARTAYGARGARYTNAARGAYGACGN 258
 QY 953 GGTTTTTGTCTGACATAATGAGCGGTGACAGCATATCAGACGCTCAGTCTCTAT 1012
 DB 259 ACNWSNTGYWSNTNCAYMGNWSNCAAYAAACNACNACNACNTAYACNTGYCAY 318
 QY 1013 ATTACTGTCTATGCCACTATGCGCAGATGACAGATCAGGTTTAAATTCGCCGATAATCCG 1072
 DB 319 ATGGAYGTNTTYTAYTATGGCNGAYGATHTTYWSNCTNAAAYATHACNACGAYCARWSN 378
 QY 1073 TCGAAGTCTGAGGATGAAGGAGGCTGAGGCTGCTTCCCTGAAAGGATATAAAGTGCATC 1132
 DB 379 GGAAAYTAYNCAARGARTGGNWSNTTYTNTGNCARGWSNMGNCARTAYAYATH 438
 QY 1133 A 1133
 DB 439 W 439

Db 379 ATGAGYNTTYCAYTYATGGGNGAYGATHTTWSNCTNAAYATHACNGAYCARWSN 438
 QY 1073 TCGAAGTCTGAGATGAAGAAGGTGAAGCTGTTCCCTGAAAGGAATAAAAGTGCATC 1132
 Db 439 GGNAAATAYWNCARGARTGYGWNSTTYTNTYNGCARGWSNMGNCARTAYAYATH 498
 QY 1133 A 1133
 Db 499 W 499
 RESULT 30
 AA10594/C
 ID AA10594 standard; DNA; 10732 BP.
 XX
 AC AA10594;
 XX 29-JUN-2000 (first entry)
 XX
 DT
 DE Gene encoding a subunit of cellulose synthase.
 XX
 KW Cellulose synthase; cellulose production; increase yield; ds.
 XX
 OS Vigna angularis.
 XX
 PN JP2000060568-A.
 XX
 PD 29-FEB-2000.
 XX
 PF 26-AUG-1998; 98JP-0239998.
 XX
 PR 26-AUG-1998; 98JP-0239998.
 XX
 PA (MIZU/) MIZUNO K.
 PA (OJIP) OJI PAPER CO.
 XX
 WPI; 2000-342371/30.
 DR P-PSDB; AAY85179.
 DR
 DR
 XX A gene encoding a cellulose synthetic equipment - for the improvement
 PT in the amount of cellulose synthesised in a plant body
 PT
 XX
 PS Claim 2; Page 14-21; 32pp; Japanese.
 XX
 CC This sequence represents a gene encoding a subunit of the cellulose
 CC synthase complex of Vigna angularis. The invention relates to subunits of
 CC cellulose synthetic equipment, that can be used to increase the amount of
 CC cellulose synthesised by a plant. The proteins and genes encoding them
 CC can also be used to improve the properties of the cellulose being
 CC produced by a plant.
 XX
 SQ Sequence 10732 BP; 3149 A; 1212 C; 2074 G; 2046 T; 2251 other;
 Query Match 3.1%; Score 36.2; DB 21; Length 10732;
 Best local similarity 15.4%; Pred. No. 2.3;
 Matches 68; Conservative 174; Mismatches 197; Indels 3; Gaps 1;
 QY 178 TCGCGGTTATCTGTGCTATTCCTTCCCTGAGCTGAGCCAGACAGACAATGATCTGTG 237
 Db 10137 DNGSRSTSTCTYAKYST 10078
 QY 238 CGCTTCTGTAATATCAACCGGTACTCAATATCTCTCTGCGCTGCGTGCATFATCC 297
 Db 10077 CYRASYDASRRYSSTYASRGTSDDCCTBSRYSCYSYDASRYANGDTBCTTTBARY- 10019
 QY 298 GGAAGCTTCCGGTCGGGATAAAATCCGAGTCCGCGGTCCATCGACACATCCC 357
 Db 10018 --RACYDAYAKRNCSTSRAYSSTCTYRCRNCYSTSTSTSTSTSTSTSTSTSTST 9961
 QY 358 CCACGGGTAAACAGCGTCCCTGTCACATCTTCTGAATGACATCAGGATCCCGCCGCT 417
 Db 9960 CTPCSRRCYSSRYSSFCNCYSYCCYTSYSTTSTSTSTSTSTSTSTSTSTSTSTST 9901

QY 418 CACTGGCGATAACGGCGACGCGGAGACTGACGCTTCCAGCCAGTACCATACCAACGCTT 477
 Db 9900 SRGYSRSGDSRGNCYNSNCTYDASTDTBYSRCCYIYSSTSDTSDTSDTSDTSDTSD 9841
 QY 478 CATTTTCCGAAGGATGACACACACACACGCGCAATCCGGTAGACCGGTAAACGCTGGGAAA 537
 Db 9840 CVTTTBSRSTSDSTSTYRCSRSDYDABSDNSTNCCYDASRTBTBSTNCACTBYDA 9781
 QY 538 GGGCACCTGCCATTAACACATCTCGCTCATCCAGGCTGTTCTGCTGCTGACGACGAC 597
 Db 9780 RCSRSTYSSRGYDANSTSRYSYSSSYSSSYSSSYSSSYSSSYSSSYSSSYSSSYSS 9721
 QY 598 GTGCTTCGTATTCTTCACGCC 619
 Db 9720 YDANCYSSDSTYBYCSRCC 9699
 RESULT 31
 ABV03347
 ID ABV03347 standard; cDNA; 650 BP.
 XX
 AC ABV03347;
 XX
 DT
 DE 13-SEP-2002 (first entry)
 XX
 DE Human prostate expression marker cDNA 3338.
 XX
 KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
 KW pharmacogenomic marker; gene; ss.
 XX
 OS Homo sapiens.
 XX
 PN WC200160860-A2.
 XX
 PD 23-AUG-2001.
 XX
 PF 20-FEB-2001; 2001WO-US05171.
 XX
 PR 17-FEB-2000; 2000US-183319P.
 PR 16-MAR-2000; 2000US-189862P.
 PR 25-MAY-2000; 2000US-207454P.
 PR 09-JUN-2000; 2000US-211314P.
 PR 18-JUL-2000; 2000US-219007P.
 PR 13-DEC-2000; 2000US-255281P.
 XX
 PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
 XX
 PI Schlegel R, Endege WO, Monahan JE;
 XX
 DR WPI; 2001-662795/76.
 XX
 PT Novel isolated nucleic acid molecule associated with cancerous state of
 PT prostate cells and correlating with presence of prostate cancer, useful
 PT for detecting presence of prostate cancer, stage of prostate cancer -
 XX
 PS Claim 1; Page 609; 11750pp; English.
 XX
 CC The invention relates to an isolated nucleic acid molecule (I) comprising
 CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
 CC specification or its complement. (I) is useful for:
 CC (a) assessing whether a patient is afflicted with prostate cancer;
 CC (b) monitoring the progression of prostate cancer in a patient;
 CC (c) assessing the efficacy of a test compound to inhibit prostate
 CC cancer in a patient;
 CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer
 CC in a patient;
 CC (e) selecting a composition for inhibiting prostate cancer in a patient;
 CC (f) assessing the prostate cell carcinogenic potential of a compound;
 CC (g) determining whether prostate cancer has metastasized in a patient;
 CC (h) assessing the aggressiveness or indolence of prostate cancer in a
 CC patient;
 CC (I) is also useful as a pharmacodynamic or pharmacogenomic marker.

XX Sequence 650 BP; 180 A; 110 C; 118 G; 199 T; 43 other;
 Query Match 3.0%; Score 36; DB 23; Length 650;
 Best Local Similarity 49.7%; Pred. No. 0.51;
 Matches 75; Conservative 0; Mismatches 76; Indels 0; Gaps 0;
 QY 14 AAAAAAGCCAAATTAATAATGCGGATCCAGCGGCTCCAGCTGAAGTAGGCGCTGT 73
 DB 479 AAAAAAGGGAATAATAATTAANAACCCGANGGTTTTTTTTTAAATAATNGNNAAT 538
 QY 74 TCTGTCGGGTATTAATGATGATGACCGTCCCGTATTAACAATGATGATAATTAATC 133
 DB 539 NGTTATTAATAATAATTTNTTCCCGCCCGCCCATTTTAACCAATNCCCTTTTCCAT 598
 QY 134 CGTTACCGGAACACCGCTGAACAAAATTCGG 164
 DB 599 GGTAAATTTTAACCAATTAANAAGGGGG 629
 RESULT 32
 ABK73466/c
 ID ABK73466 standard; DNA; 1427 BP.
 AC ABK73466;
 XX
 XX
 DT 13-AUG-2002 (first entry)
 XX
 DE Bacillus licheniformis genomic sequence tag (GST) #757.
 KW
 KW Differential gene expression; genomic sequenced tag; GST;
 KW altered culture condition; environmental stress;
 KW physiological provocation; ds.
 XX
 XX Bacillus licheniformis.
 XX WO200229113-A2.
 XX
 PD 11-APR-2002.
 XX
 XX 05-OCT-2001; 2001WO-US31437.
 XX
 XX 06-OCT-2000; 2000US-0680598.
 PR 27-MAR-2001; 2001US-279526P.
 PA (NOVO) NOVOZYMES BIOFTECH INC.
 PA (NOVO) NOVOZYMES AS.
 XX
 XX Berka R, Clausen IG;
 XX WPI; 2002-416684/44.
 DR
 XX Monitoring differential expression of several genes in first Bacillus
 PT cell relative to expression of same genes in one or more second
 PT Bacillus cells, by using substrate containing Bacillus genomic
 PT sequenced tag array -
 XX
 PS Claim 4; SEQ ID NO 757; 200pp; English.
 XX
 CC The invention describes a method of monitoring differential expression of
 CC genes in a first Bacillus cell relative to expression of the genes in
 CC other Bacillus cells, comprising hybridising labelled nucleic acid probes
 CC isolated from Bacillus cells to a substrate containing array of Bacillus
 CC genomic sequenced tags (GST), examining the array, and determining
 CC relative gene expression by an observed hybridisation reporter signal of
 CC a spot in the array. The method is useful for measuring the expression of
 CC genes in a first Bacillus cell relative to expression of the same genes
 CC in one or more second Bacillus cells. The method is useful for monitoring
 CC global expression of several genes from a Bacillus cell, discovering new
 CC genes, identifying possible functions of unknown open reading frames and
 CC monitoring gene copy number variation and stability. Monitoring changes
 CC in expression of genes may be used to provide a representation of the way
 CC in which Bacillus cells adapt to changes in culture conditions,

CC environmental stress or other physiological provocation. Extensive
 CC follow-up characterisation is unnecessary, when one spot on an array
 CC equals one gene or one open reading frame, since sequence information is
 CC available. This sequence represents a genomic sequence tag (GST) used in
 CC the method of the invention.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at
 CC ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 1427 BP; 378 A; 305 C; 379 G; 365 T; 0 other;
 Query Match 3.0%; Score 36; DB 24; Length 1427;
 Best Local Similarity 46.0%; Pred. No. 0.81;
 Matches 160; Conservative 0; Mismatches 185; Indels 3; Gaps 1;
 QY 398 ATCAGGGATCCCGCGCTGCTCACTGGCGATAAGCGGCACCGGAGACTGACGCTTCAGC 457
 DB 1215 ATCATGAAGCCCGCTGTTTCCCTGACACCGGCACCGCGCTGAGGCGAATCAG 1156
 QY 458 CAGTACCATACCAACGCTTCATTTTCCGAAGCATGACACCACTGGCAATCCGGTA 517
 DB 1155 CTGTCCAAGACCGCAGCGCTCGAATTTTGAAGCATCAGGAAAGATCGCTCGCGGCATA 1096
 QY 518 GACCGGTAACGCTGGGAAAGGCGACCTGCCATTAACATATCCGCTCATTCACAGGTG 577
 DB 1095 AATCTTGGCGGTGAGCGGTTTCATGAAACCGCATATACGCTTGACCTGTCGGAAGAAGC 1036
 QY 578 TTCTGTCTGCTGACGAGAGCGTCTGCTGTTATTTTCAGCGCGCGCCACACGAGCCA 637
 DB 1035 ATGCTCCATATACCGGAAGTAATTTTCAAAATTCGGCTTCTCCGTTTCTTAAACGATCAG 976
 QY 638 GCGAAATGATTTCCCTTCCATCTTCAGCTGATACATACAGCAGCATATAATTCATGTC 697
 DB 975 CTGCATGTCTTCTTCTTCA---ACAGTTTCATCAATACCGCTTGATCAATCGAAGCC 919
 QY 698 TTTTTCGGGAGTAGCATCCCACTGACGATAAGCGGAACATTTGTC 745
 DB 918 TTTTTCGTTGTCAGCGCTTGCCACCATGCTGATGAGCGGAATGCTTCTC 871
 RESULT 33
 AAH29949/c
 ID AAH29949 standard; DNA; 1859 BP.
 XX
 XX AAH29949;
 XX
 DT 27-JUL-2001 (first entry)
 XX
 DE S cerevisiae apoptosis associated coding sequence YKR040C.
 XX
 KW Yeast; fungus; apoptosis; infection; proliferative disease;
 KW vaccine; autoimmune disease; ischaemia; neurodegeneration; ds.
 XX
 OS Saccharomyces cerevisiae.
 XX
 PN WO200102550-A2.
 XX
 PD 11-JAN-2001.
 XX
 PF 03-JUL-2000; 2000WO-BE00077.
 XX
 PR 01-JUL-1999; 99EP-0870141.
 XX
 PA (JANC) JANSSEN PHARM NV.
 XX
 PI Contreras RH, De Backer MD, Luyten WHML, Malcorps IKL;
 PI Nelissen BJM, Reekmans RJ;
 XX
 XX WPI; 2001-367042/38.
 DR P-PSDB; AAG70913.
 XX
 PT Yeast and fungal nucleic acids encoding proteins involved in a pathway

PT leading to programmed cell death, useful for treating proliferative
 PT disorders, yeast and fungal infections, or for preventing apoptosis in
 XX certain diseases
 PS Claim 1; Fig 1; 218pp; English.
 XX The present invention provides the protein and coding sequences of a
 CC number of apoptosis associated proteins from the yeast *Saccharomyces*
 CC *cerevisiae* and the fungus *Candida albicans*. These can be used to identify
 CC treatments for fungal and yeast infections, for proliferative diseases
 CC and for apoptosis related diseases such as autoimmune diseases, ischaemia
 CC and neurodegeneration. The present sequence is one of the *S. cerevisiae*
 CC coding sequences of the invention.

XX Sequence 1859 BP; 663 A; 297 C; 358 G; 541 T; 0 other;

Query Match 3.0%; Score 36; DB 22; Length 1859;
 Best Local Similarity 56.9%; Pred. No. 0.95;
 Matches 66; Conservative 0; Mismatches 50; Indels 0; Gaps 0;
 QY 375 CCTGTGACATCTCTGAATGACATCAGGATCCGCCGCTCTCACTGGCGATAACGGGC 434
 DB 1459 CTGTGATCTGTTGGTAAGACTTCGGGTATGCTTCCCACTTTTGTAGTGACACATAT 1400
 QY 435 ACGCCGGAGACTGACGTTTCAGGCAGTACCATACCAACGCTTCATTTCCGAAGG 490
 DB 1399 AACCCACATGATGCAGCTTCCACAATAACTGTACCAACGCTCTGTATATGAAGG 1344

RESULT 34

AA565293/c
 ID AA565293 standard; cDNA; 426 BP.

XX AC AA565293;

XX DT 13-FEB-2002 (first entry)

XX DE DNA encoding novel human diagnostic protein #1097.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX OS Homo sapiens.

XX PN WO200175067-A2.

XX PD 11-OCT-2001.

XX PF 30-MAR-2001; 2001WO-US08631.

XX PR 31-MAR-2000; 2000US-0540217.

XX PR 23-AUG-2000; 2000US-0649167.

XX PA (HYSE-) HYSEQ INC.

XX PI Drmanac RT, Liu C, Tang YT;

XX DR WPI; 2001-639362/73.

XX DR P-PSDB; ABG01106.

XX New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity

XX Claim 1; SEQ ID No 1097; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques

CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. AA564197-AA594564 represent novel human
 CC diagnostic coding sequences of the invention.

CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 426 BP; 113 A; 105 C; 113 G; 95 T; 0 other;

Query Match 3.0%; Score 35.8; DB 23; Length 426;
 Best Local Similarity 65.8%; Pred. No. 0.47;
 Matches 52; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

QY 525 AACGCTGGGAAAGGCGACCTGCCATTACACATCTCCGCTCATCCAGGTGTTCTGTC 584
 DB 394 ATTCTTAGACAGACGACCTTCCAGGACACAGCTCCCCCCTCACTGCTGGACCTTC 335
 QY 585 TGCTGACGACGACGCTGCTT 603
 DB 334 CTCGTGGGAAGACTGGCTT 316

RESULT 35

ABV48872
 ID ABV48872 standard; cDNA; 488 BP.

XX AC ABV48872;

XX DT 17-SEP-2002 (first entry)

XX Human prostate expression marker cDNA 48863.

XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
 KW pharmacogenomic marker; gene; ss.

XX OS Homo sapiens.

XX PN WO200160860-A2.

XX PD 23-AUG-2001.

XX PF 20-FEB-2001; 2001WO-US05171.

XX PR 17-FEB-2000; 2000US-183319P.

XX PR 16-MAR-2000; 2000US-189862P.

XX PR 25-MAY-2000; 2000US-207454P.

XX PR 09-JUN-2000; 2000US-211314P.

XX PR 18-JUL-2000; 2000US-219007P.

XX PR 13-DEC-2000; 2000US-255281P.

XX PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX PI Schlegel R, Endege WO, Monahan JE;

XX DR WPI; 2001-662795/76.

XX Novel isolated nucleic acid molecule associated with cancerous state of
 PT prostate cells and correlating with presence of prostate cancer, useful
 PT for detecting presence of prostate cancer, stage of prostate cancer

XX Claim 1; Page 9567; 11750pp; English.

XX The invention relates to an isolated nucleic acid molecule (I) comprising
 CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the

XX DE Human nervous system related polynucleotide SEQ ID NO 11358.
XX DE
XX KW Human; nootropic; neuroprotective; cytostatic; dermatological; virucide;
KW KW immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnery;
KW KW antiparkinsonian; antisickling; antianaemic; antiarthritic; cancer;
KW KW antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;
KW KW antiallergic; antidiabetic; antitumor; anticonvulsant; antifungal;
KW KW antiparasitic; cardiant; immune disorder; cardiovascular disorder;
KW KW neurological disease; infection; nephrotropic; gene therapy; vaccine; ds.
XX OS Homo sapiens.
XX OS
XX PN WO200159063-A2.
XX PD
XX PD 16-AUG-2001.
XX PD
XX PD 17-JAN-2001; 2001WO-US01334.
XX PF
XX PF
XX PF 31-JAN-2000; 2000US-0179065.
XX PR 04-FEB-2000; 2000US-0180828.
XX PR 24-FEB-2000; 2000US-0184664.
XX PR 02-MAR-2000; 2000US-0186350.
XX PR 16-MAR-2000; 2000US-0189874.
XX PR 17-MAR-2000; 2000US-0190076.
XX PR 19-APR-2000; 2000US-0198123.
XX PR 19-MAY-2000; 2000US-0205515.
XX PR 07-JUN-2000; 2000US-0209467.
XX PR 28-JUN-2000; 2000US-0214886.
XX PR 30-JUN-2000; 2000US-0215135.
XX PR 07-JUL-2000; 2000US-0216647.
XX PR 07-JUL-2000; 2000US-0216880.
XX PR 11-JUL-2000; 2000US-0217487.
XX PR 11-JUL-2000; 2000US-0217496.
XX PR 14-JUL-2000; 2000US-0218290.
XX PR 26-JUL-2000; 2000US-0220963.
XX PR 26-JUL-2000; 2000US-0220964.
XX PR 14-AUG-2000; 2000US-0224518.
XX PR 14-AUG-2000; 2000US-0224519.
XX PR 14-AUG-2000; 2000US-0225213.
XX PR 14-AUG-2000; 2000US-0225214.
XX PR 14-AUG-2000; 2000US-0225266.
XX PR 14-AUG-2000; 2000US-0225267.
XX PR 14-AUG-2000; 2000US-0225268.
XX PR 14-AUG-2000; 2000US-0225270.
XX PR 14-AUG-2000; 2000US-0225447.
XX PR 14-AUG-2000; 2000US-0225757.
XX PR 14-AUG-2000; 2000US-0225758.
XX PR 14-AUG-2000; 2000US-0225759.
XX PR 18-AUG-2000; 2000US-0226279.
XX PR 22-AUG-2000; 2000US-0226681.
XX PR 22-AUG-2000; 2000US-0226868.
XX PR 23-AUG-2000; 2000US-0227182.
XX PR 23-AUG-2000; 2000US-0227009.
XX PR 30-AUG-2000; 2000US-0228924.
XX PR 01-SEP-2000; 2000US-0229287.
XX PR 01-SEP-2000; 2000US-0229343.
XX PR 01-SEP-2000; 2000US-0229344.
XX PR 01-SEP-2000; 2000US-0229345.
XX PR 05-SEP-2000; 2000US-0229509.
XX PR 05-SEP-2000; 2000US-0229513.
XX PR 08-SEP-2000; 2000US-0230437.
XX PR 08-SEP-2000; 2000US-0230438.
XX PR 08-SEP-2000; 2000US-0231242.
XX PR 08-SEP-2000; 2000US-0231243.
XX PR 08-SEP-2000; 2000US-0231244.
XX PR 08-SEP-2000; 2000US-0231413.
XX PR 08-SEP-2000; 2000US-0231414.
XX PR 08-SEP-2000; 2000US-0232080.
XX PR 08-SEP-2000; 2000US-0232081.
XX PR 12-SEP-2000; 2000US-0231968.
XX PR 14-SEP-2000; 2000US-0232397.
XX PR 14-SEP-2000; 2000US-0232398.
XX PR 14-SEP-2000; 2000US-0232399.
XX PR 14-SEP-2000; 2000US-0232400.
XX PR 14-SEP-2000; 2000US-0232401.
XX PR 14-SEP-2000; 2000US-0233063.
XX PR 14-SEP-2000; 2000US-0233064.
XX PR 14-SEP-2000; 2000US-0233065.
XX PR 21-SEP-2000; 2000US-0234223.
XX PR 21-SEP-2000; 2000US-0234274.
XX PR 25-SEP-2000; 2000US-0234997.
XX PR 25-SEP-2000; 2000US-0234998.
XX PR 26-SEP-2000; 2000US-0235484.
XX PR 27-SEP-2000; 2000US-0235834.
XX PR 27-SEP-2000; 2000US-0235836.
XX PR 29-SEP-2000; 2000US-0236327.
XX PR 29-SEP-2000; 2000US-0236367.
XX PR 29-SEP-2000; 2000US-0236368.
XX PR 29-SEP-2000; 2000US-0236369.
XX PR 29-SEP-2000; 2000US-0236370.
XX PR 02-OCT-2000; 2000US-0236802.
XX PR 02-OCT-2000; 2000US-0237037.
XX PR 02-OCT-2000; 2000US-0237038.
XX PR 02-OCT-2000; 2000US-0237039.
XX PR 02-OCT-2000; 2000US-0237040.
XX PR 13-OCT-2000; 2000US-0239335.
XX PR 13-OCT-2000; 2000US-0239337.
XX PR 20-OCT-2000; 2000US-0240960.
XX PR 20-OCT-2000; 2000US-0241785.
XX PR 20-OCT-2000; 2000US-0241786.
XX PR 20-OCT-2000; 2000US-0241787.
XX PR 20-OCT-2000; 2000US-0241808.
XX PR 20-OCT-2000; 2000US-0241809.
XX PR 20-OCT-2000; 2000US-0241826.
XX PR 20-OCT-2000; 2000US-0242221.
XX PR 01-NOV-2000; 2000US-0244617.
XX PR 08-NOV-2000; 2000US-0246474.
XX PR 08-NOV-2000; 2000US-0246475.
XX PR 08-NOV-2000; 2000US-0246476.
XX PR 08-NOV-2000; 2000US-0246477.
XX PR 08-NOV-2000; 2000US-0246478.
XX PR 08-NOV-2000; 2000US-0246523.
XX PR 08-NOV-2000; 2000US-0246524.
XX PR 08-NOV-2000; 2000US-0246525.
XX PR 08-NOV-2000; 2000US-0246526.
XX PR 08-NOV-2000; 2000US-0246527.
XX PR 08-NOV-2000; 2000US-0246528.
XX PR 08-NOV-2000; 2000US-0246532.
XX PR 08-NOV-2000; 2000US-0246609.
XX PR 08-NOV-2000; 2000US-0246610.
XX PR 08-NOV-2000; 2000US-0246611.
XX PR 08-NOV-2000; 2000US-0246613.
XX PR 17-NOV-2000; 2000US-0249207.
XX PR 17-NOV-2000; 2000US-0249208.
XX PR 17-NOV-2000; 2000US-0249209.
XX PR 17-NOV-2000; 2000US-0249210.
XX PR 17-NOV-2000; 2000US-0249211.
XX PR 17-NOV-2000; 2000US-0249212.
XX PR 17-NOV-2000; 2000US-0249213.
XX PR 17-NOV-2000; 2000US-0249214.
XX PR 17-NOV-2000; 2000US-0249215.
XX PR 17-NOV-2000; 2000US-0249216.
XX PR 17-NOV-2000; 2000US-0249217.
XX PR 17-NOV-2000; 2000US-0249218.
XX PR 17-NOV-2000; 2000US-0249244.
XX PR 17-NOV-2000; 2000US-0249245.
XX PR 17-NOV-2000; 2000US-0249264.
XX PR 17-NOV-2000; 2000US-0249265.
XX PR 17-NOV-2000; 2000US-0249297.
XX PR 17-NOV-2000; 2000US-0249299.
XX PR 01-DEC-2000; 2000US-0249300.
XX PR 01-DEC-2000; 2000US-0250391.
XX PR 05-DEC-2000; 2000US-0251160.
XX PR 05-DEC-2000; 2000US-0251030.
XX PR 05-DEC-2000; 2000US-0251988.

PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254037.
PR 05-JAN-2001; 2001US-0259678.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
PI
XX
XX WPI; 2001-541565/60.
DR
XX
XX Nucleic acids encoding 3224 human nervous system antigen polypeptides,
PT useful for preventing, diagnosing and/or treating nervous system
PT cancers and metastases -
XX
XX Disclosure; SEQ ID NO 11358; 1701pp + Sequence Listing; English.
XX
XX The invention relates to novel genes (ABAI1004-ABA21534) and proteins
CC (ABAI14678-ABAI18001) useful for preventing, treating or ameliorating
CC medical conditions e.g. by protein or gene therapy. The genes are
CC isolated from a range of human tissues disclosed in the specification.
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful
CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
CC and ovarian cancer and other cancers of the adrenal gland, bone, bone
CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune
CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative
CC colitis; (c) cardiovascular disorders such as myocardial ischaemias;
CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
CC and parasitic infections.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 5666 BP; 1450 A; 1247 C; 1332 G; 1637 T; 0 other;
SQ

Query Match 3.0%; Score 35.8; DB 22; Length 5666;
Best Local Similarity 65.8%; Pred. No. 2.1;
Matches 52; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

QY 525 AACGCTGGGAAGGACCTGCCATTACACATCTCCGTCATTCCAGGTGTTCTGTC 584
Db 4528 AATTCTTAGCAGACGACCTTCAGGAAACAGCTCCGCCACTCACTGGTGACCTTC 4469

QY 585 TGCTGACGACGACGTGCTT 603
Db 4468 CTCTGGGAAGACTGCTT 4450

RESULT 38
AAK75576/c
ID AAK75576 standard; DNA; 5666 BP.
XX
XX
AC AAK75576;
XX
XX 07-NOV-2001 (first entry)
DT
XX Human immune/haematopoietic antigen genomic sequence SEQ ID NO:30388.
DE Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
XX cytostatic; gene therapy; vaccine; metastasis; ds.
KW
XX Homo sapiens.
XX
XX WO200157182-A2.
PN
XX

PD 09-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US01354.
XX
XX 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232081.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.

29-SEP-2000; 2000US-0236368.
29-SEP-2000; 2000US-0236369.
29-SEP-2000; 2000US-0236370.
02-OCT-2000; 2000US-0236802.
02-OCT-2000; 2000US-0237037.
02-OCT-2000; 2000US-0237038.
02-OCT-2000; 2000US-0237039.
02-OCT-2000; 2000US-0237040.
13-OCT-2000; 2000US-0239935.
13-OCT-2000; 2000US-0239937.
20-OCT-2000; 2000US-0240960.
20-OCT-2000; 2000US-0241221.
20-OCT-2000; 2000US-0241785.
20-OCT-2000; 2000US-0241786.
20-OCT-2000; 2000US-0241787.
20-OCT-2000; 2000US-0241808.
20-OCT-2000; 2000US-0241809.
20-OCT-2000; 2000US-0241826.
01-NOV-2000; 2000US-0244617.
08-NOV-2000; 2000US-0246474.
08-NOV-2000; 2000US-0246475.
08-NOV-2000; 2000US-0246476.
08-NOV-2000; 2000US-0246477.
08-NOV-2000; 2000US-0246478.
08-NOV-2000; 2000US-0246523.
08-NOV-2000; 2000US-0246524.
08-NOV-2000; 2000US-0246525.
08-NOV-2000; 2000US-0246526.
08-NOV-2000; 2000US-0246527.
08-NOV-2000; 2000US-0246528.
08-NOV-2000; 2000US-0246532.
08-NOV-2000; 2000US-0246532.
08-NOV-2000; 2000US-0246609.
08-NOV-2000; 2000US-0246610.
08-NOV-2000; 2000US-0246611.
08-NOV-2000; 2000US-0246613.
17-NOV-2000; 2000US-0249207.
17-NOV-2000; 2000US-0249208.
17-NOV-2000; 2000US-0249209.
17-NOV-2000; 2000US-0249210.
17-NOV-2000; 2000US-0249211.
17-NOV-2000; 2000US-0249212.
17-NOV-2000; 2000US-0249213.
17-NOV-2000; 2000US-0249214.
17-NOV-2000; 2000US-0249215.
17-NOV-2000; 2000US-0249216.
17-NOV-2000; 2000US-0249217.
17-NOV-2000; 2000US-0249218.
17-NOV-2000; 2000US-0249244.
17-NOV-2000; 2000US-0249245.
17-NOV-2000; 2000US-0249264.
17-NOV-2000; 2000US-0249265.
17-NOV-2000; 2000US-0249297.
17-NOV-2000; 2000US-0249299.
01-DEC-2000; 2000US-0250160.
01-DEC-2000; 2000US-0250391.
05-DEC-2000; 2000US-0251030.
05-DEC-2000; 2000US-0251988.
05-DEC-2000; 2000US-0256719.
06-DEC-2000; 2000US-0256719.
08-DEC-2000; 2000US-0251479.
08-DEC-2000; 2000US-0251856.
08-DEC-2000; 2000US-0251868.
08-DEC-2000; 2000US-0251869.
08-DEC-2000; 2000US-0251989.
08-DEC-2000; 2000US-0251990.
11-DEC-2000; 2000US-0254097.
05-JAN-2001; 2001US-0259678.
(HUMA-) HUMAN GENOME SCI INC.
Rosen CA, Barash SC, Ruben SM;
WPI; 2001-483426/52.

XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT useful for preventing, diagnosing and/or treating cancers and
PT metastasis -
XX
PS Disclosure; SEQ ID NO 30388; 3071pp + Sequence Listing; English.
XX
CC AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
CC amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patients own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting
CC the nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/hematopoietic-related diseases, especially
CC cancers and cancer metastases of hematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/hematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169
CC represent sequences used in the exemplification of the present invention.
XX
SQ Sequence 5666 BP; 1450 A; 1247 C; 1332 G; 1637 T; 0 other;

Query Match 3.0%; Score 35.8; DB 22; Length 5666;
Best Local Similarity 65.8%; Pred. No. 2.1;
Matches 52; Conservative 0; Mismatches 27; Indels 0; Gaps 0;
QY 525 AACGCTGGGAAGGCGACCTGCCATTACACATCTCCGTCATTCGCCAGGTGTTGTC 584
DB 4528 AATTCCTAGACGAGACACCTTCCAGGAACACAGCTCCGCCACTCAGTGGGACCTTC 4469
QY 585 TGCTGACGACGACGTGCTT 603
DB 4468 CTCCTGGGAGACTGGCTT 4450

RESULT 39
AAD06413
ID AAD06413 standard; DNA; 1614 BP.
XX
AC AAD06413;
XX
DT 10-AUG-2001 (first entry)
XX
DE Human DNAX cytokine receptor subunit 3.2 reverse translational DNA.
XX
KW Human; immunomodulator; DNAX cytokine receptor subunit 3.2; DCRS3.2;
KW therapy; immunological disorder; drug screening; cell development;
XX chromosome 16p12; ds.
XX
OS Homo sapiens.
XX
XX WO200136467-A2.
XX
XX 25-MAY-2001.
XX
XX 16-NOV-2000; 2000WO-US31363.
XX
XX 18-NOV-1999; 99US-0443060.
PR 13-DEC-1999; 99US-0170320.
XX
XX (SCHE) SCHERING CORP.
XX
PI Gorman DM;
XX
XX WPI; 2001-343800/36.
XX
XX New mammalian receptor proteins related to cytokine receptors, useful
PT for regulating cell development and for diagnosis and treatment of

PT immunological disorders
PS Disclosure; Page 18; 124pp; English.
XX
CC The present sequence is human DNAX cytokine receptor subunit 3.2
CC (DCRS3.2) reverse translational DNA. DCRS3 gene is located on
CC chromosome 16p12. Cytokine receptors, fragments and antibodies are
CC useful for treating immunological disorders. DCRS3 (50R), DCRS4 (cytor)
CC or fragments are useful in drug screening to identify compounds having
CC binding affinity to the receptor subunit. Modulators of DCRS are useful
CC for modulating the physiology or development of a cell or tissue culture
CC cells. A purified DCRS is useful as a reagent to detect antibodies
CC generated in response to the presence of elevated levels of expression,
CC or immunological disorders which lead to production of antibody to the
CC endogenous receptor. Cytokine receptor sequences are useful as probes
CC for detecting levels of the cytokine receptor in patients suspected of
CC having an immunological disorder. Antibodies have therapeutic value, are
CC useful as potent antagonist, in detecting or quantifying ligands, for
CC isolating DCRS proteins and peptides, to screen expression libraries for
CC particular expression products, to raise anti-idiotypic antibodies and
CC for detecting or diagnosing various immunological conditions related to
CC expression of the protein or cells which express the protein.
XX
SQ Sequence 1614 BP; 206 A; 189 C; 302 G; 202 T; 715 other;
Query Match 3.0%; Score 35.6; DB 22; Length 1614;
Best Local Similarity 30.3%; Pred. No. 1.2;
Matches 73; Conservative 34; Mismatches 134; Indels 0; Gaps 0;
QY 883 GTGATAGGACAAACATGAATACACAGATTAATTCAGGGAAGGTTCTGGTCTTAC 942
Db 117 GTGAAATTTTCAC 176
QY 943 GGGTGTAGTGGTTTGTCTGACATAGTGAAGCGTGACAGCATATCAGACGGGTCA 1002
Db 177 RGATGACGAC 236
QY 1003 GTCTGCTATATCTACTGTCGACATGACGACATGACGACATGACGACATGACG 1062
Db 237 YACNTGTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 296
QY 1063 CGATATCCGTCGAGTCTGAGGATGAGGATGAGGATGAGGATGAGGATGAGGAT 1122
Db 297 NGATCARWSNGNAAT 356
QY 1123 A 1123
Db 357 R 357
RESULT 40
AAZ61263/c
ID AAZ61263 standard; DNA; 2163 BP.
XX
AC AAZ61263;
XX
DT 30-MAY-2000 (first entry)
XX
DE DNA encoding a sucrose phosphate synthase-like enzyme.
XX
KW Sucrose biosynthesis; sucrose phosphate synthase-like enzyme;
KW SPS-like enzyme; glucosyl donor substrate; sucrose; sucrose phosphate;
KW tobacco; rice; carbon partitioning; growth response;
KW CDP-glucose-dependent sucrose-6'-phosphate synthase;
KW sucrose-6'-phosphate; nucleoside diphosphate glucose;
KW fructose-6-phosphate; ss.
XX
OS Synchocystis sp.
XX
FH Location/Qualifiers
FT 1..2163
FT /*tag= a
FT /product= "sucrose phosphate synthase-like enzyme"

XX
PN WO200003006-A1.
XX
PD 20-JAN-2000.
XX
PF 08-JUL-1999; 99WO-AU00557.
XX
PR 08-JUL-1999; 98AU-0004578.
XX
PA (CSIR) COMMONWEALTH SCI IND & RES ORG.
XX
PI Furbank R, Lunn J;
XX
PIE WPI; 2000-182220/16.
XX
DR P-PSDB; AAY69235.
XX
PT Novel nucleic acids encoding sucrose phosphate synthase-like enzymes,
PT useful for altering the level of sucrose or sucrose phosphates and
PT altering carbon partitioning in plants
XX
PS Claim 16; Page 91-95; 99pp; English.
XX
CC The present sequence encodes a sucrose biosynthesis polypeptide which
CC has a sucrose phosphate synthase-like (SPS-like) activity. The SPS-like
CC polypeptide does not utilize UDP-glucose as a sole glucosyl donor
CC substrate. Probes and amplification primers based on the SPS-like
CC polynucleotide can be used to isolate or detect nucleic acid molecules
CC which encode a protein with SPS-like activity. The SPS-like
CC polynucleotide can be used to alter the level sucrose or sucrose
CC phosphates in a plant cell, tissue, or organ or a higher plant,
CC especially tobacco or rice. The SPS-like polynucleotide can also be
CC used to alter the carbon partitioning, development and/or growth
CC response of a plant, increase the level of UDP-, GDP-, ADP- or
CC CDP-glucose-dependent sucrose-6'-phosphate synthesis, and to produce
CC sucrose-6'-phosphate from a nucleoside diphosphate glucose other than
CC UDP-glucose in a plant cell. Recombinant SPS-like activity expressed
CC in E. coli shows a broader specificity than the higher plant SPS
CC enzymes, in particular utilising ADP-, CDP-, GDP- or UDP-glucose as a
CC glucosyl donor to fructose-6-phosphate.
XX
SQ Sequence 2163 BP; 557 A; 517 C; 543 G; 546 T; 0 other;
Query Match 3.0%; Score 35.6; DB 21; Length 2163;
Best Local Similarity 54.6%; Pred. No. 1.4;
Matches 71; Conservative 0; Mismatches 59; Indels 0; Gaps 0;
QY 375 CCTGTACATCTCTGTAATGATCAGGATCCCGCCGCTCTCACTGGCGATAACGGGC 434
Db 1169 CCATTCGACAAATTTTGTGATATATCCACGGGCCCCCATCTCGTGCCCAATGGGC 1110
QY 435 ACGCCGGAGACTGACGCTTCAGCAGTACCATACCAACGCTTCATTTCCGAAGGCATG 494
Db 1109 ACACACACAGCGCGCTTCATCAATCAAGTTAAACCAAGGTTCCGTCACCAAGGATG 1050
QY 495 ACCACACACAC 504
Db 1049 ATAAATATCTC 1040
RESULT 41
AAS59545/c
ID AAS59545 standard; DNA; 61313 BP.
XX
AC AAS59545;
XX
DT 13-FEB-2002 (first entry)
XX
DE Propionibacterium acnes immunogenic protein encoding DNA #40.
XX
KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KW dermatological; osteopathic; neuroprotectant; ds.

Db 847 GCCCTCCCTGATGCTTGGAGCGACACACACGCTGCAAGCTG 805

RESULT 43

AAC21988/c
ID AAC21988 standard; cDNA; 355 BP.

XX AC AAC21988;

XX 06-OCT-2000 (first entry)

XX Human secreted protein 5' EST, SEQ ID NO: 26063.

XX Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW gene therapy; chromosome mapping; ss.

XX Homo sapiens.

XX EP1033401-A2.

XX 06-SEP-2000.

XX 21-FEB-2000; 2000EP-0200610.

XX 26-FEB-1999; 99US-0122487.

XX (GEST) GENSET.

XX Dumas Milne Edwards J, Duclert A, Giordano J;

XX WPI; 2000-500381/45.

XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT diagnostic, forensic, gene therapy and chromosome mapping procedures -

PS Claim 1; SEQ ID 26063; 71pp + CD-ROM; English.

XX The present sequence is one of a large number of 5' ESTs derived from
CC mRNAs encoding secreted proteins. No ORF has yet been conclusively
CC identified within the present sequence. The 5' ESTs were prepared from
CC total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
CC sequences usually correspond mainly to the 3' untranslated region (UTR)
CC of the mRNA because they are often obtained from oligo-dT primed cDNA
CC libraries. Such ESTs are not well suited for isolating cDNA sequences
CC derived from the 5' ends of mRNAs and even in those cases where longer
CC cDNA sequences have been obtained, the full 5' UTR is rarely included.
CC 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be
CC used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used
CC in diagnostic, forensic, gene therapy and chromosome mapping procedures.
CC They are used to obtain upstream regulatory sequences and to design
CC expression and secretion vectors.

SQ Sequence 355 BP; 90 A; 57 C; 68 G; 119 T; 11 other;

Query Match 2.9%; Score 34.6; DB 21; Length 355;
Best Local Similarity 48.4%; Pred. NO. 1;
Matches 77; Conservative 6; Mismatches 76; Indels 0; Gaps 0;

QY 139 CCGGAACCCGCTGAACAAATTCGGCTGAAAGAGGATCCGCCGTTATCTGTTCATT 198

DB 329 CTGAATAACAAGAGCTCTCAATTGAGGGGGAAGAGCCCTCCAVACTGACATCTARRAT 270

QY 199 TCCCTTTAGCTGACAGACACAATGATCTGTGCGCTTCTGTTATATCAAAACC 258

DB 269 TCCTCTTAGAGACACACAGTACTKCCAAACAACATCTTTCTKGTTTAATWTGAATACC 210

QY 259 GGTACTCAATATCTTCTGCGGCTGGCTGCATCATCC 297

DB 209 CSRACCTGGTCATATCTTTGAAGACTGCGCTTCATATAC 171

RESULT 44

AAK81290
ID AAK81290 standard; DNA; 21410 BP.
XX AC AAK81290;
XX 07-NOV-2001 (first entry)
XX Human Immune/haematopoietic antigen genomic sequence SEQ ID NO:36102.
KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis; ds.
XX Homo sapiens.
XX WO200157182-A2.
XX 09-AUG-2001.
XX 17-JAN-2001; 2001WO-US01354.
XX 31-JAN-2000; 2000US-0179065.
XX 04-FEB-2000; 2000US-0180628.
XX 24-FEB-2000; 2000US-0184664.
XX 02-MAR-2000; 2000US-0186350.
XX 16-MAR-2000; 2000US-0189874.
XX 17-MAR-2000; 2000US-0190076.
XX 18-APR-2000; 2000US-0198123.
XX 19-MAY-2000; 2000US-0205515.
XX 07-JUN-2000; 2000US-0209467.
XX 28-JUN-2000; 2000US-0214886.
XX 30-JUN-2000; 2000US-0215135.
XX 07-JUL-2000; 2000US-0216847.
XX 07-JUL-2000; 2000US-0216880.
XX 11-JUL-2000; 2000US-0217487.
XX 11-JUL-2000; 2000US-0217496.
XX 14-JUL-2000; 2000US-0218290.
XX 26-JUL-2000; 2000US-0220963.
XX 26-JUL-2000; 2000US-0220964.
XX 14-AUG-2000; 2000US-0224518.
XX 14-AUG-2000; 2000US-0224519.
XX 14-AUG-2000; 2000US-0225213.
XX 14-AUG-2000; 2000US-0225214.
XX 14-AUG-2000; 2000US-0225266.
XX 14-AUG-2000; 2000US-0225267.
XX 14-AUG-2000; 2000US-0225268.
XX 14-AUG-2000; 2000US-0225270.
XX 14-AUG-2000; 2000US-0225447.
XX 14-AUG-2000; 2000US-0225757.
XX 14-AUG-2000; 2000US-0225758.
XX 14-AUG-2000; 2000US-0225759.
XX 18-AUG-2000; 2000US-0226279.
XX 22-AUG-2000; 2000US-0226681.
XX 22-AUG-2000; 2000US-0226868.
XX 22-AUG-2000; 2000US-0227182.
XX 23-AUG-2000; 2000US-0227009.
XX 30-AUG-2000; 2000US-0228924.
XX 01-SEP-2000; 2000US-0229287.
XX 01-SEP-2000; 2000US-0229343.
XX 01-SEP-2000; 2000US-0229344.
XX 01-SEP-2000; 2000US-0229345.
XX 05-SEP-2000; 2000US-0229509.
XX 05-SEP-2000; 2000US-0229513.
XX 06-SEP-2000; 2000US-0230437.
XX 06-SEP-2000; 2000US-0230438.
XX 08-SEP-2000; 2000US-0231242.
XX 08-SEP-2000; 2000US-0231243.
XX 08-SEP-2000; 2000US-0231244.
XX 08-SEP-2000; 2000US-0231413.
XX 08-SEP-2000; 2000US-0231414.
XX 08-SEP-2000; 2000US-0232080.
XX 08-SEP-2000; 2000US-0232081.
XX 12-SEP-2000; 2000US-0231968.
XX 14-SEP-2000; 2000US-0232397.


```

XX OS Homo sapiens.
XX FH Key
XX FH CDS
XX FT Location/Qualifiers
XX FT /tag= a "Human phosphatase protein"
XX FT /product= (3114, A)
XX FT /tag= b
XX FT /standard_name= "Single nucleotide polymorphism (SNP)"
XX FT /note= "Polymorphic site 1"
XX FT /tag= c
XX FT /tag= c (4004, A)
XX FT /standard_name= "Single nucleotide polymorphism (SNP)"
XX FT /note= "Polymorphic site 2"
XX FT /tag= d
XX FT /tag= d (4514, G)
XX FT /standard_name= "Single nucleotide polymorphism (SNP)"
XX FT /note= "Polymorphic site 3"
XX FT /tag= e
XX FT /tag= e (7570, G)
XX FT /standard_name= "Single nucleotide polymorphism (SNP)"
XX FT /note= "Polymorphic site 4"
XX FT /tag= f
XX FT /tag= f (11672, G)
XX FT /standard_name= "Single nucleotide polymorphism (SNP)"
XX FT /note= "Polymorphic site 5"
XX FT /tag= g
XX FT /tag= g (11897, C)
XX FT /standard_name= "Single nucleotide polymorphism (SNP)"
XX FT /note= "Polymorphic site 6"
XX FT /tag= h
XX FT /tag= h (14523, C)
XX FT /standard_name= "Single nucleotide polymorphism (SNP)"
XX FT /note= "Polymorphic site 7"
XX FT /tag= i
XX FT /tag= i (16586, T)
XX FT /standard_name= "Single nucleotide polymorphism (SNP)"
XX FT /note= "Polymorphic site 8"
XX FT /tag= j
XX FT /tag= j (16644, C)
XX FT /standard_name= "Single nucleotide polymorphism (SNP)"
XX FT /note= "Polymorphic site 9"
XX FT /tag= k
XX FT /tag= k (17969, G)
XX FT /standard_name= "Single nucleotide polymorphism (SNP)"
XX FT /note= "Polymorphic site 10"
XX FT /tag= l
XX FT /tag= l (18117, T)
XX FT /standard_name= "Single nucleotide polymorphism (SNP)"
XX FT /note= "Polymorphic site 11"
XX FT /tag= m
XX FT /tag= m (18518, A)
XX FT /standard_name= "Single nucleotide polymorphism (SNP)"
XX FT /note= "Polymorphic site 12"
XX FT /tag= n
XX FT /tag= n (19364..19420)
XX FT /tag= o
XX FT /tag= o (19421..34110)
XX FT /tag= p
XX FT /tag= p (19882, A)
XX FT /standard_name= "Single nucleotide polymorphism (SNP)"
XX FT /note= "Polymorphic site 13"
XX FT /tag= q
XX FT /tag= q (20988, -)
XX FT /standard_name= "Single nucleotide polymorphism (SNP)"
XX FT /note= "Polymorphic site 14"
XX FT /tag= r
XX FT /tag= r (20999, T)
XX FT /standard_name= "Single nucleotide polymorphism (SNP)"
XX FT /note= "Polymorphic site 15"
XX FT /tag= s
XX FT /tag= s (21465, G)

```

```

FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT /note= "Polymorphic site 16"
FT /tag= t
FT /tag= t (21625, T)
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT /note= "Polymorphic site 17"
FT /tag= u
FT /tag= u (26291, T)
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT /note= "Polymorphic site 18"
FT /tag= v
FT /tag= v (28012, C)
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT /note= "Polymorphic site 19"
FT /tag= w
FT /tag= w (28030, G)
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT /note= "Polymorphic site 20"
FT /tag= x
FT /tag= x (33671, C)
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT /note= "Polymorphic site 21"
FT /tag= y
FT /tag= y (34111..34143)
FT /tag= z
FT /tag= z (34144..35683)
FT /tag= aa
FT /tag= aa (35684..35737)
FT /tag= ab
FT /tag= ab (35738..39940)
FT /tag= ac
FT /tag= ac (37703, G)
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT /note= "Polymorphic site 22"
FT /tag= ad
FT /tag= ad (39263, G)
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT /note= "Polymorphic site 23; This position is given as 39269 in the specification"
FT /tag= ae
FT /tag= ae (39941..40038)
FT /tag= af
FT /tag= af (40039..45810)
FT /tag= ag
FT /tag= ag (45811..45871)
FT /tag= ah
FT /tag= ah (45872..46578)
FT /tag= ai
FT /tag= ai (46579..46615)
FT /tag= aj
FT /tag= aj (46616..47002)
FT /tag= ak
FT /tag= ak (47003..47042)
FT /tag= al
FT /tag= al (47043..47133)
FT /tag= am
FT /tag= am (47134..47184)
FT /tag= an
FT /tag= an (47185..48943)
FT /tag= ao
FT /tag= ao (48944..49016)
FT /tag= ap
FT /tag= ap (49017..57568)
FT /tag= aq
FT /tag= aq (57569..57602)
FT /tag= ar
FT /tag= ar (57603..57761)
FT /tag= as
FT /tag= as (57762..59835)

```

WO200160992-A2.

23-AUG-2001.

PD

PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 29-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 07-JUN-1999; 99US-0137502.
PR 08-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 25-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151303.
PR 27-AUG-1999; 99US-0151080.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.

RESULT 48	
AAK60217	
ID	AAK60217 standard; cDNA; 333 BP.
XX	
XX	
AC	
XX	AAK60217;
XX	
DT	06-NOV-2001 (first entry)
XX	
DE	Human immune/haematopoietic antigen encoding cDNA SEQ ID NO:5277.
XX	
KW	Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW	cystostatic; gene therapy; vaccine; metastasis; ss.
XX	
OS	Homo sapiens.
XX	
PN	WO200157182-A2.
XX	

PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

PI Schlegel R, Endege WO, Monahan JE;

XX WPI; 2001-662795/76.

XX Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer -

XX Claim 1; Page 9566-9567; 11750pp; English.

XX The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV0010-ABV62213) of the
CC specification or its complement. (I) is useful for:

CC (a) assessing whether a patient is afflicted with prostate cancer;

CC (b) monitoring the progression of prostate cancer in a patient;

CC (c) assessing the efficacy of a test compound to inhibit prostate

CC cancer in a patient;

CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer

CC in a patient;

CC (e) selecting a composition for inhibiting prostate cancer in a patient;

CC (f) assessing the prostate cell carcinogenic potential of a compound;

CC (g) determining whether prostate cancer has metastasized in a patient;

CC (h) assessing the aggressiveness or indolence of prostate cancer in a

CC patient;

CC (I) is also useful as a pharmacodynamic or pharmacogenomic marker.

XX Sequence 474 BP; 140 A; 114 C; 118 G; 102 T; 0 other;

Query Match 2.9%; Score 33.8; DB 23; Length 474;

Best Local Similarity 55.68; Pred. No. 2.3;

Matches 65; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

QY 7 GAGATGGAAAAAGCCAAATAAATAATGCCCCATCCAGCGCGCTCCAGCTGAAGA 66

DB 238 GAAATAAAAAAAGAAAAAAGGCGCGCCAGCGGGTGAGCTCCAATTC 297

QY 67 GGCGTGTCTGCGGTATTAATGCAATGACCGTCCCGGTATTTAAACAATGTGA 123

DB 298 GCCCTATAGTAGTCGTATTACGCGCGCTCACTGGCGTGTGTTTACAGCGTGTA 354

RESULT 50

AAQ21621/c

ID AAQ21621 standard; DNA; 1647 BP.

XX AAQ21621;

XX 29-JUN-1992 (first entry)

XX Human Tryptophan-2,3-dioxygenase coding sequence HTO3.

XX Tourette Syndrome; TDO2; kyneurenine formation; autism; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

XX CDS 34..1344

XX /*tag= a

XX CDS 34..1008

XX /*tag= b

XX WO9202637-A.

XX 20-FEB-1992.

XX 02-AUG-1991; 91WO-US05441.

XX 14-JUN-1991; 91US-0715660.

XX 03-AUG-1990; 90US-0562596.

XX (CITY) CITY OF HOPE.

XX Comings DE;

XX WPI; 1992-080090/10.

XX P-PSDB; AAR21547.

XX Diagnosis of Tourette syndrome and associated disorders - by
PT detecting mutation in tryptophan oxygenase-indole:amine oxygenase
PT genes and prods.

XX Disclosure; Fig 5A; 45pp; English.

XX Clone HTO3 was isolated from a human liver mRNA-derived cDNA
CC library screened with a rat liver tryptophan oxygenase cDNA probe
CC (see AAQ21622).

CC The specification includes two different sequences which are

CC both described as "HTO3" (see also AAQ21620). They are identical

CC through the first 741 nucleotides of AAQ21620 before the much longer

CC sequence AAQ21621 begins to diverge. The human TDO2 gene sequence,

CC including regulatory sequences, is also given in the specification,

CC as is a genomic TDO2 sequence which includes the introns (see AAQ21623

CC and AAQ21624, respectively). The sequence of the open reading frame in

CC these sequences differs from AAQ21621 (and AAQ21620) by the deletion of

CC a single adenine residue (i.e. from the run of 8 A's at position

CC 713-720 of AAQ21621). Yet another Figure, which illustrates the

CC degree of homology between HTO3 and the rat TDO2 sequence, also

CC lacks the 8th Adenine, suggesting that the "extra" A in both

CC AAQ21620 and AAQ21621 represents a typographical error. The patent

CC specification does not contain any discussion of the apparent

CC differences and includes only the minimum description of each of

CC the sequences. Consequently, the significance of an in frame

CC termination codon within the longest CDS of HTO3 is not known.

XX Sequence 1647 BP; 556 A; 292 C; 327 G; 472 T; 0 other;

Query Match 2.9%;

Best Local Similarity 53.48; Pred. No. 4.7; Length 1647;

Matches 71; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

QY 1013 ATTACTGTCTGCGCACTATGGCAGATGACCCAGATCAGGTTTAAATCCCGATATCCG 1072

DB 345 ATCTCTCACATGGCCATTCTGAAAGATCTCTCGAACACAGATCCCACTCCACAGAGATTG 286

QY 1073 TCGAAGTCTGAGGATGGAAGGAGCTGCTTCTTGAAGGATAAAGTGACATC 1132

DB 285 CTTAAACCAGAGTTTCATAGCTTGATGAGTTATGATAAAAGATGTTTCATCATGATTTT 226

QY 1133 ATGCCCTCTTTT 1145

DB 225 ATTTCCCTTTGTT 213

Search completed: June 18, 2003, 02:09:39

Job time : 223.603 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 18, 2003, 01:20:59 ; Search time 1249.12 Seconds
(without alignments)
15312.305 Million cell updates/sec

Title: US-09-674-277-2
Perfect score: 1181
Sequence: 1 ctgcaggagatgagaaaaa.....ttttactttttctctgcag 1181

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 150 summaries

Database :

- EST:*
- 1: em_estba:*
 - 2: em_esthum:*
 - 3: em_estin:*
 - 4: em_estnu:*
 - 5: em_estov:*
 - 6: em_estpl:*
 - 7: em_estro:*
 - 8: em_hic:*
 - 9: gb_est1:*
 - 10: gb_est2:*
 - 11: gb_hic:*
 - 12: gb_est3:*
 - 13: gb_est4:*
 - 14: gb_est5:*
 - 15: em_estfun:*
 - 16: em_estom:*
 - 17: gb_gss:*
 - 18: em_gss_hum:*
 - 19: em_gss_inv:*
 - 20: em_gss_pln:*
 - 21: em_gss_vrt:*
 - 22: em_gss_fun:*
 - 23: em_gss_mam:*
 - 24: em_gss_mus:*
 - 25: em_gss_other:*
 - 26: em_gss_pro:*
 - 27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	48.8	4.1	755	17	AQ989708
2	46.6	3.9	615	10	AV909088
3	46.6	3.9	629	13	BJ264853
4	46.6	3.9	669	10	AV910971
5	46.6	3.9	693	10	AV945422
6	45.8	3.9	647	13	BJ312175

495	14	BQ763870	BQ763870	EBR003_SQ
714	9	AU100720	AU100720	BOH0324TF
866	17	BH468052	BH468052	Drosophila
1101	17	CNS0106X	CNS0106X	AW042265
533	11	AW042265	AW042265	614026D01
736	11	AY112421	AY112421	Zea mays
645	14	BQ087067	BQ087067	Cri_9_106
461	17	BH778732	BH778732	BH778732 f2mb013f0
769	17	BH388561	BH388561	AG-ND-130
38.6	12	BG330039	BG330039	602428912
38	17	BI744669	BI744669	XK89H09.Y
332	13	BG605346	BG605346	WHE2331.F
479	13	BJ292385	BJ292385	BJ292385
545	14	C20457	C20457	Rice
756	17	CNS02D50	CNS02D50	Tetraodon
649	10	AV784670	AV784670	AV784670
501	12	BG840188	BG840188	WEST8-G12
540	12	BF586903	BF586903	FM1_31_HO
672	13	BI968888	BI968888	GM830006B
499	13	BI498039	BI498039	sag15c08.
516	13	BM143281	BM143281	sa142b12.
609	10	AV836868	AV836868	AV836868
849	17	CNS031JD	CNS031JD	Tetraodon
974	14	BQ878757	BQ878757	AGENCOURT
948	17	CNS017IM	CNS017IM	Drosophila
320	10	AW453731	AW453731	zeh10506
831	12	BG565743	BG565743	602589078
582	10	AV987761	AV987761	AV987761
1101	17	CNS00LT2	CNS00LT2	Drosophila
861	12	BG338455	BG338455	602587161
432	10	AW509231	AW509231	SH92H11.Y
445	14	BQ611687	BQ611687	sap63g01.
446	10	AW459606	AW459606	sh89c12.Y
478	9	AI496532	AI496532	sh13a03.Y
507	10	AW620859	AW620859	sh47e11.Y
512	17	AQ785148	AQ785148	HS_3159_A
514	17	AQ662456	AQ662456	HS_5455_A
523	10	BE440931	BE440931	sp49g12.Y
527	10	AW279073	AW279073	sg07d10.Y
540	13	BM092322	BM092322	ash11e10.Y
542	13	BI427174	BI427174	eah77c01.
566	13	BI427022	BI427022	sag10d03.
581	14	BQ630772	BQ630772	sap31a10.
584	13	BI787127	BI787127	sag71f10.
585	10	AW397142	AW397142	sg67d08.Y
586	14	BQ629541	BQ629541	saq03d02.
636	10	AW164393	AW164393	se71h10.Y
821	12	BG618482	BG618482	602645674
1101	17	CNS014TP	CNS014TP	Drosophila
319	9	AI229512	AI229512	EST226207
482	12	BF565804	BF565804	UT-R-C4-a
557	14	BM860712	BM860712	fy53f12.X
248	10	AV358623	AV358623	AV358623
254	17	AQ078429	AQ078429	CIT-HSP-2
372	12	BF404709	BF404709	UT-R-CAL-
412	10	BB699898	BB699898	BB699898
434	13	BB699994	BB699994	BB699994
438	13	BJ449621	BJ449621	BJ449621
520	10	AV906979	AV906979	AV906979
523	13	BJ450987	BJ450987	BJ450987
545	14	BP017181	BP017181	BP017181
600	13	BJ458496	BJ458496	BJ458496
650	10	BE279974	BE279974	601156988
406	17	AQ184201	AQ184201	HS_3221_A
646	14	BQ562752	BQ562752	4077G06-
663	10	BB662428	BB662428	BB662428
972	14	BQ715820	BQ715820	AGENCOURT
3321	11	AK014488	AK014488	Mus muscu
576	17	AQ231338	AQ231338	HS_2053_A
616	9	AL635769	AL635769	AL635769
712	12	BF273334	BF273334	GA_EB001
793	13	BI821211	BI821211	603034921
925	17	CNS0091P	CNS0091P	Drosophila

Tel: 81-559-81-6856			
Fax: 81-559-81-6855			
Email: tshini@genes.nig.ac.jp.			
Location/Qualifiers			
1. .629	/organism="Triticum aestivum"		
	/cultivar="Chinese Spring"		
	/db_xref="taxon:4565"		
	/clone="wh9f24"		
	/tissue_type="spike at heading date"		
	/dev_stage="Feekes' scale 10.5"		
	/note="Vector: Lambda Uni-ZAP XR, excised phagemid; Site_1: EcoRI; Site_2: XhoI; Plants were grown under hydroponic conditions at UC Davis, salt stressed for 12 h, and for 7 days, then dissected and frozen (Akhunov in J. Dvorak Lab). Total RNA was prepared from sheath tissue, equal quantities of RNA were pooled from the two samples, polyA was purified from the pooled RNA, a cDNA library was made, and the cDNA clones were in vivo excised to give pBluescript phagemids in the TJ Close lab at the University of California, Riverside (Akhunov, Chin, Choi, Close, Fenton, Kianian, Otto, Simons, Zhang). Plasmid DNA preparations and DNA sequencing were performed in the OD Anderson lab (all other authors)."		
BASE COUNT	146 a 191 c 138 g 151 t 3 others		
ORIGIN			
Query Match	3.9%;	Score 46.6;	DB 13; Length 629;
Best Local Similarity	56.1%;	Pred. No. 0.043;	
Matches	88; Conservative	0; Mismatches	69; Indels 0; Gaps 0;
Qy	371 CGTCCCTGTGCACATCTCTGAATGACATCGAGGATCCGCGCGTCTCACTGGCGATAAC	430	
Db			
Qy	329 CTTCTCTCTGATCCTCAGGTATATGTTCAGGTATGCTCCAGCTCGAGCGACCAC	388	
Db			
Qy	431 GGGACGCGGAGCATGACCGCTTCAGCAGTACCATACCAACGCTTCATTTCCGAAGG	490	
Db			
Qy	389 CGGGATCCGGATGACATTCCTCCACACAAACCCAGCGTCTCGGACTCAGAAGG	448	
Db			
Qy	491 CATGACCAACACACTGCGATCCGGTAGACCGGTAAAC	527	
Db			
	449 CATCACAACACATCCCGACTGGCATAGCGCTGTGAC	485	
RESULT 4			
AV910971			
LOCUS	AV910971 K. Sato unpublished cDNA library, linear EST 18-JAN-2002		
DEFINITION	vegetative stage leaves Hordeum vulgare subsp. vulgare cv. Akashinriki baak1122 3', mRNA sequence.		
ACCESSION	AV910971		
VERSION	AV910971.1 GI:18206747		
KEYWORDS	EST.		
SOURCE	Hordeum vulgare subsp. vulgare.		
ORGANISM	Hordeum vulgare subsp. vulgare		
REFERENCE	Sato,K., Saisho,D. and Takeda,K.		
AUTHORS	Barley EST sequencing project in NIG and Okayama Univ		
TITLE	Unpublished (2002)		
JOURNAL	Contact: Tadasu Shin-i		
COMMENT	Center For Genetic Resource Information		
	National Institute of Genetics		
	1111 Yata, Mishima, Shizuoka 411-8540, Japan		
	Tel: 81-559-81-6856		
	Fax: 81-559-81-6855		
	Email: tshini@genes.nig.ac.jp.		
FEATURES	Location/Qualifiers		
source	1. .669		
	/organism="Hordeum vulgare subsp. vulgare"		


```

RESULT 7
BQ763870/c
LOCUS      BQ763870             495 bp     mRNA          linear       EST 26-JUL-2002
DEFINITION EBR03_SQ005_K13_R root, 3 week, waterlogged, cv Optic, EBr003
            Hordeum vulgare cDNA clone EBr03_SQ005_K13 5', mRNA sequence.

ACCESSION  BQ763870
VERSION     BQ763870.1    GI:21972342
KEYWORDS    EST.
SOURCE      BQ763870.1    GI:21972342
ORGANISM    Hordeum vulgare.
            Hordeum vulgare.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
; Triticeae; Hordeum.
1 (bases 1 to 495)
Hedley,P., Liu,H., Caldwell,D., McCallum,N., Mudie,S., Cardle,L.,
Ramsay,L., Machray,G., Marshall,D.F.M. and Waugh,R.
Development of Barley Transcriptome Resources
Unpublished (2001)
Contact: Waugh R, Marshall DF
Genome Dynamics/Computational Biology
Scottish Crop Research Institute
Invergowrie, Dundee, DD2 5DA, Scotland, UK
Tel.: 00 44 1382 562731
Fax: 00 44 1382 562426
Email: est@scri.sari.ac.uk.

FEATURES             Location/Qualifiers
     source           1..495
                     /organism="Hordeum vulgare"
                     /cultivar="Optic"
                     /db_xref="taxon:4513"
                     /clone="EBR03_SQ005_K13"
                     /clone_lib="root, 3 week, waterlogged, cv Optic, EBro03"
                     /tissue_type="root"
                     /dev_stage="3 week"
                     /lab_host="DH10B"
     note="Vector: pSPORt1; Site_1: Sal I; Site_2: Not I;
Non-normalised library, directionally cloned into pSPORt1.
Derived from roots of 3 week old waterlogged barley
plants. Developed as part of the barley transcriptome
resources of BSRC/SERAD funded cereal IGF (Investigating
Gene function) project."
BASE COUNT   53 a   176 c   189 g   77 t
ORIGIN
Query Match              3.8%; Score 45; DB 14; Length 495;
Best Local Similarity    49.8%; Pred.No. 0.12; Indels 0; Gaps 0;
Matches 114; Conservative 0; Mismatches 115;

Qy      404  GATCCGCCCGCTCTCACTGGCGATAACGCGCAGCGGAGACTGTTCAGCGTTCAGCCAGTAC 463
         ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
Db      261  GAAGTGACGCAGCGCTGCGAGGAGCGCGCGCGCATGAGCGCCACGCGCGGGC 202

Qy      464  CATACCAAAGCGTTTCATTFTTCGAAGAGCATGACCAACACACTGGCAATCGGTAGACCGG 523
         ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
Db      201  CACGCCGAGCACCGCATGTCGCCGCGGATGAGATGATCCCGCGCGCCCACACGCGTCGT 142

Qy      524  TAACGTGGAAAAGGCGACCTGCCATTACACATCTCCGCTCATTTCCAGGTTGTCTGT 583
         ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
Db      141  CGACGTTTGTAACACCGCGCCGAGGACCGGCTTGC CGGGAGGCGCCCCCGCTCCGTG 82

Qy      584  CTGCTGACGCAGCGTGTCTGTATTCTTCACGCCCGCGGCCACCCACG 632
         ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
Db      81  CTCGGGAGCAGNGCGCCTCTCCGCTTCGCTGCGCGCGGTGGCCACG 33


RESULT 8
AU100720/c
LOCUS      AU100720             714 bp     mRNA          linear       EST 03-APR-2002
DEFINITION AU100720 Rice callus Oryza sativa (japonica cultivar-group) CDNA
            clone Cl2279, mRNA sequence.

ACCESSION  AU100720
```



```

QY 444 ACTGACGCTTCAGCCAGTACCATACCAACGCTTCATTTCCTCCGAGGCGATGACCACCACA 503
Db 241 GCATGGACTCAGCACTACTTGCCTGGAGGGCTCGGATTCGAAGCGATTGCAATATACA 182
QY 504 CTGGCAATCCGGTAGACCGGTAA 526
Db 181 TCCGCACTGGCTACGCTTGTA 159

RESULT 12
AY112421
LOCUS Zea mays CL8980_1 mRNA sequence. 736 bp mRNA linear HTC 26-MAY-2002
DEFINITION Zea mays CL8980_1 mRNA sequence.
ACCESSION AY112421
VERSION AY112421.1 GI:21217011
KEYWORDS HTC.
SOURCE Zea mays.
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 736)
Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S.,
Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.
Maize Mapping Project/DuPont Consensus Sequences for Design of
Overgo Probes
Unpublished (2002)
REFERENCE 2 (bases 1 to 736)
AUTHORS Coe,E.C.
TITLE Direct Submission
JOURNAL Submitted (25-APR-2002) Maize Mapping Project, University of
Missouri, Columbia, MO 65211, USA
FEATURES
source
1..736
/organism="Zea mays"
/db_xref="MaizeDB:633031"
/db_xref="taxon:4577"
/clone="CL8980_1"
/clone_lib="Maize Mapping Project/DuPont Consensus
Library"
/notes="this sequence is part of a project of EST
assemblies resulting from the application of public
contigs to seed DuPont contigs; this resource was
assembled by DuPont as part of a collaboration for the
overgo addressing of BACS in conjunction with the Maize
Mapping Project."
BASE COUNT 171 a 198 c 154 g 182 t 31 others
ORIGIN
Query Match 3.4%; Score 40.6; DB 11; Length 736;
Best Local Similarity 55.2%; Pred. No. 2.2;
Matches 79; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

QY 384 TTCTTCTGATGACATCAGCATCCCGCGCTCTACTGGCGATACAGGCGACGGCGGAG 443
Db 438 TCCTTGGGTATATATACAGTATCCCTCCAGCAGCAGCAGCAGCTGGACTCCAGAA 497
QY 444 ACTGACGCTTCAGCCAGTACCATACCAACGCTTCATTTCCTCCGAGGCGATGACCACCACA 503
Db 498 GCCATGGACTCCAGCACTACTTGCCTGGAGGGTCTCGGATTCCTGAAGGCGATTGCAATACA 557
QY 504 CTGGCAATCCGGTAGACCGGTAA 526
Db 558 TCCGCACTGGCTACGCTTGTA 580

RESULT 13
BQ087067/c
LOCUS Zea mays CL8980_1 mRNA sequence. 645 bp mRNA linear EST 05-APR-2002
DEFINITION Zea mays CL8980_1 mRNA sequence.
ACCESSION BQ087067
VERSION BQ087067.1 GI:20046268

```

```

KEYWORDS Ceratopteris richardii.
SOURCE Ceratopteris richardii
ORGANISM Eukaryota; Viridiplantae; Embryophyta; Tracheophyta;
Filicophyta; Filicopsida; Filicales; Pteridaceae; Ceratopteris.
1 (bases 1 to 645)
Chatterjee,A., San Miguel,P., Stout,S.C., Banks,J. and Roux,S.J.
Expressed sequence tags of cDNA clones from a C. richardii library
Unpublished (2000)
CONTACT: Roux SJ
Section of Molecular Cell and Developmental Biology
University of Texas
Biology Building, Room 16, Austin, TX 78712, USA
Tel: 512 471 4238
Fax: 512 232 3402
Email: sroux@uts.cc.utexas.edu
Plate: Cri_9 row: I column: 06
Seq primer: SP6.
FEATURES
source
1..645
/organism="Ceratopteris richardii"
/cultivar="Brogn"
/db_xref="taxon:49495"
/clone="Cri_9_I06"
/clone_lib="Ceratopteris Spore Library"
/tissue_type="Gametophyte"
/cell_type="Spore"
/dev_stage="20 hours after germination initiation"
/notes="Vector: pCMVSPORT6; EST sequence from cDNA library.
cDNA library constructed from mRNA isolated from C.
richardii spores that had developed for 20 hours after
their germination had been initiated by white light."
BASE COUNT 159 a 128 c 186 g 171 t 1 others
ORIGIN
Query Match 3.3%; Score 39.4; DB 14; Length 645;
Best Local Similarity 53.8%; Pred. No. 4.7;
Matches 82; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

QY 374 CCTGTGCATCTTCTGATGACATCAGGATCCGCGCTCTCACTGGCGATGACGGG 433
Db 535 CCCCTCCTGATGGCTGGGATATATCAGGAATGCCCTGCTCTTCTGCTCCCAATGGG 476
QY 434 CAGCGCGGAGACTGACGCTTCACCGAGTACCATTACCAACGCTTCATTTCCGAAGCAT 493
Db 475 TACCCCTGATGCTATAGCCTCAAGAACACACAGCAAGAGTTTCTGATTTCAGAGGGTGT 416
QY 494 GACCAACACATCGGCAATCCGCTAGACCGGTAA 526
Db 415 CACAAAAACATCCCACTTGCATAGGCTTGTA 383

RESULT 14
BQ087067/c
LOCUS Zea mays CL8980_1 mRNA sequence. 461 bp DNA linear GSS 28-MAR-2002
DEFINITION Zea mays CL8980_1 mRNA sequence.
ACCESSION BQ087067
VERSION BQ087067.1 GI:19781391
KEYWORDS GSS.
SOURCE Zea mays.
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 461)
Budiman,M.A., Freese,R.G., Bedell,J.A., Nunberg,A.N. and Lakey,N.D.
Generalthresh methylation filtered genomic sequences from maize
Unpublished (2002)
CONTACT: Bedell JA
Orion Genomics, LLC
4041 Forest Park Ave, St. Louis, MO 63108, USA
Tel: 314 615 6979

```

Fax: 314 615 5975
Email: jbedell@oriongenomics.com
Plate: f2mb013f051 row: a column: 07
Seq primer: M13 forward
Class: shotgun
High quality sequence stop: 461.
Location/Qualifiers
1. 461
/organism="Zea mays"
/cultivar="M017"
/db_xref="taxon:4577"
/clone_lib="f2mb013f051a07"
/note="Organ: leaf; Vector: pBCSK(-); Site_1: HindIII; DNA prepared from purified nuclei was randomly sheared, end-repaired, size fractionated to enrich for the 0.5 to 5 kb fraction, ligated into HindIII-digested pBCSK(-) vector and electroporated into E. coli cells."
BASE COUNT 122 a 98 c 120 g 121 t
ORIGIN
Query Match 3.3%; Score 39.2; DB 17; Length 461;
Best Local Similarity 53.2%; Pred. No. 5.1;
Matches 83; Conservative 0; Mismatches 73; Indels 0; Gaps 0;
QY 371 GGTCCCTGTACATCTTCTGAATGACATCAGGATCCCGCGTCTCACTGGCGATAAC 430
DB 399 CTTACCCCTCTGTGCTTCGGTATATATCAGGTATCCCTCCAGCAGCAGCAGCAGC 340
QY 431 GGGCAGCGGAGACTGACGCTTCAGCCAGTACCATACCAACAGCTTCATTTCCGAAGG 490
DB 339 TGGGATCCAGACCATGAGTCTCCAGCACTATTCGCGAGGTCTCGGATTCGTAAGG 280
QY 491 CATGACACACACATCGCATCCGGTAGACGGGTAA 526
DB 279 CATTGCAATACATCTCCACTGGCGTACGCTTGTA 244
RESULT 15
BH388561 769 bp DNA linear GSS 11-DEC-2001
LOCUS AG-ND-130K10.TF ND-TAM Anopheles gambiae genomic clone AG-ND-130K10
DEFINITION , DNA sequence.
ACCESSION BH388561
VERSION BH388561.1 GI:17334702
KEYWORDS GSS
SOURCE African malaria mosquito.
ORGANISM Anopheles gambiae
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoides; Anopheles.
REFERENCE 1 (bases 1 to 769)
Shetty J., Malek J., Koo H., Collins F., Gardner M. and Loftus B.J. Direct Submission of BAC-end sequences from Anopheles gambiae Unpublished (2001)
Other_GSSs: AG-ND-130K10.TF
Contact: Brendan J Loftus
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0208
Fax: 301 838 3543
Email: b.loftus@tigr.org
This clone is from an A. gambiae BAC library (ND-TAM) provided by F.H. Collins and sequenced by The Institute for Genomic Research (TIGR). The BAC library was generated from A. gambiae pBst strain DNA. All DNA was extracted from newly hatched first instar larvae to minimize the inclusion of DNA from microorganisms that inhabit the gut. The DNA is derived from mixed sexes of larvae. The BAC library was constructed at Texas A&M University BAC Center University, College Station, Texas 77843-2123, USA using a HindIII Seq primer: M13 For

Class: BAC ends.
Location/Qualifiers
1. 769
/organism="Anopheles gambiae"
/strain="PES1"
/db_xref="taxon:7165"
/clone_lib="AG-ND-130K10"
/note="Vector: pECBAC1; Site_1: HindIII"
BASE COUNT 212 a 157 c 143 g 257 t
ORIGIN
Query Match 3.3%; Score 38.6; DB 17; Length 769;
Best Local Similarity 52.1%; Pred. No. 8;
Matches 86; Conservative 0; Mismatches 79; Indels 0; Gaps 0;
QY 345 GCAGACACATCCCCACGGTAACAGCTCCCTGTCACATCTTCTGAATGACATCAGG 404
DB 137 GCATCTACATCCCAACATCTGCAATAATCCGGTTACTCCGTGTTTGTACTCCGGT 196
QY 405 ATCCGCGCGTCTCACTGGCGATAACGGCGAGACTGAGCTTCAGCCAGTACC 464
DB 197 ATTCGCGCGCATTAATCTACTGATTACCGGAGTTCGGCAGCATCTGCTCCAGAGTCA 256
QY 465 ATACCAAGCGTTTCATTTCCGAAGCATGACCACACACTGGCA 509
DB 257 AGACCGAGCTTCTCTGTTGTGACGGAAGCATATAATACATCCGCA 301
RESULT 16
BG330039 926 bp mRNA linear EST 27-FEB-2001
LOCUS 602428912F1 NIH_MGC_16 Homo sapiens cDNA clone IMAGE:455814 5', mRNA sequence.
DEFINITION mRNA sequence.
ACCESSION BG330039
VERSION BG330039.1 GI:13136477
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 926)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgraphs@mail.nih.gov
Tissue procurement: ATCC
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: image.llnl.gov Plate: LICM1264 row: h column: 03
High quality sequence stop: 667.
Location/Qualifiers
1. 926
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:455814"
/clone_lib="NIH_MGC_16"
/tissue_type="retinoblastoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: eye; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; CDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."
BASE COUNT 129 a 287 c 264 g 246 t
ORIGIN

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percormorpha; Tetraodontiformes; Tetraodontidae; Tetraodon.

REFERENCE AUTHORS

1 (bases 1 to 756)
Roest-Crolius, H., Jaillon, O., Dasilva, C., Bouneau, L., Fisher, C., Bernot, A., Fizames, C., Wincker, P., Brottier, P., Quetier, F., Saurin, W. and Weissenbach, J.
Human gene number estimate provided by genome wide analysis using Tetraodon nigroviridis DNA sequence

JOURNAL REFERENCE AUTHORS

2 (bases 1 to 756)
Roest-Crolius, H., Jaillon, O., Dasilva, C., Fizames, C., Fisher, C., Bouneau, L., Billault, A., Quetier, F., Saurin, W., Bernot, A. and Weissenbach, J.
Characterization and repeat analysis of the compact genome of the freshwater pufferfish Tetraodon nigroviridis

JOURNAL REFERENCE AUTHORS

3 (bases 1 to 756)
Direct Submission
Submitted (12-APR-2000)

JOURNAL COMMENT

This sequence is a single read and was generated as part of a large scale clone-end sequencing project of the Tetraodon nigroviridis genome. For more information, please take a look at <http://www.genoscope.cns.fr/Tetraodon>.

FEATURES source

Location/Qualifiers
1..756
/organism="Tetraodon nigroviridis"
/db_xref="taxon:99883"
/clone="256P09"
/clone_lib="G"
/note="Genoscope sequence ID : C0AG256CH05SP1-end : PUC-ori"

BASE COUNT 217 a 156 c 204 g 174 t 5 others
ORIGIN

Query Match 3.2%; Score 37.4; DB 17; Length 756;
Best Local Similarity 48.8%; Pred. No. 17;
Matches 101; Conservative 0; Mismatches 106; Indels 0; Gaps 0;
QY 967 CAATAGTGAAGGGTGCACATATCAGACGGCTCAGTCTGTCTATTATTATTCATGTC 1026
DB 333 CATTATTGAAGTGGCTGTGGACCTCTGTAGTCTCAATCAATCTTATTACCTGTGGG 274
QY 1027 CACTATFGCGAGATGACCATCAGGTTTAAATTCGCCGATAATCCGTAAGTCTGAGGA 1086
DB 273 AAATATGGCTGGCAATCACCGTACAAAGTAATGATACATAAACCATACAGAGAT 214
QY 1087 TGGAGGAGGTGAAGCTGTTCCTGAAGGATATAAGTGACATCATGCCCTTTTTC 1146
DB 213 GATTCAACAGACACCACCATCGACCTGGCAGGAAGAAGTAATATTTCCGCTGTTTTC 154
QY 1147 TGGCTTCGGGAGCAATTTTACTTTTTT 1173
DB 153 CACCTACAGTCAGACATCTATTGT 127

RESULT 22 AV784670 LOCUS

DEFINITION AV784670 RAFL6 Arabidopsis thaliana cDNA clone RAFL06-07-E03 3', mRNA sequence.
VERSION AV784670
KEYWORDS EST.
SOURCE thale cress.

ORGANISM REFERENCE AUTHORS

Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
1 (bases 1 to 649)
Seki, M., Narusaka, M., Ishida, J., Kamiya, A., Satou, M., Nakajima, M.,

TITLE JOURNAL COMMENT

Large scale analysis of Arabidopsis full-length cDNA (2002b)
Contact: Motoaki Seki
Plant Functional Genomics Research Group
RIKEN Genomic Sciences Center
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
Tel: 81-298-36-4359
Fax: 81-298-36-9060
Email: msekic@rct.riken.go.jp

An Arabidopsis full-length cDNA library was constructed essentially as reported previously (Seki et al., 1998). This clone is in a modified pluscript vector as a SstI/XhoI insert. Please visit our web site (http://www.gsc.riken.go.jp/e/plant/index_e.html) for further details.

FEATURES source

Location/Qualifiers
1..649
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"
/clone="RAFL06-07-E03"
/clone_lib="RAFL6"
/dev_stage="plants at various developmental stages from germination to mature seeds"
/lab_host="DH10B"
/note="Site 1: SstI; Site 2: XhoI; subjected to dehydration (1, 2, 5, 10, 24 hr) and cold (1, 2, 5, 10, 24 hr) treatments"

BASE COUNT 182 a 173 c 102 g 192 t
ORIGIN

Query Match 3.1%; Score 37.2; DB 10; Length 649;
Best Local Similarity 52.6%; Pred. No. 20;
Matches 81; Conservative 0; Mismatches 73; Indels 0; Gaps 0;
QY 373 TCCTGTGCATCTTCTGTGATGACATCAGGATCCGGCCGCTCTCCTGCGCATACGG 432
DB 453 TTCCCTCCCTGATCTTCAGGGATTATATCAGGATTCCCGGCCGCGCAGCAGCA 512
QY 433 GCACGCGGGAGCTGAGCTTCAGCCAGTACCATCAACAGCTTCATTTTCCGAAGGCA 492
DB 513 GAAGTCTGAGACATGCTTCAAGAACCAAGGCCAAGTGTCTCGACTCTGTGGCA 572
QY 493 TGACCAACACATCGCAATCCGGTAGACCGGTAA 526
DB 573 TCACAAACACATCTCCACTTGCCTAAGCTTGTC 606

RESULT 23 BG840188 LOCUS

DEFINITION BG840188 501 bp mRNA linear EST 29-MAY-2001
ACCESSION BG840188
VERSION BG840188.2 GI:14243001
KEYWORDS EST.
SOURCE Zea mays.
ORGANISM Zea mays.

REFERENCE AUTHORS TITLE JOURNAL COMMENT

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 501)
Qiu, F., Cui, F., Guo, L., Ashlock, D.A., Wen, T.J. and Schnable, P.S.
Expressed Sequence Tags from B73 Maize Seedlings and Silks
Unpublished (2001)
On May 25, 2001 this sequence version replaced gi:14206510.
Contact: Patrick S. Schnable
Iowa State University
G405 Agronomy, Iowa State University, Ames, IA 50011-1010, USA
Tel: 515-294-0975
Fax: 515-294-2299

Email: schnable@iastate.edu

PCR Primers
 FORWARD: T7-1 (AA TAC GAC TCA CTA TAG)
 BACKWARD: T3 (ATT AAC CCT CAC TAA AG)
 Seq primer: primer T3 (ATT AAC CCT CAC TAA AG).

FEATURES
 Source
 Location/Qualifiers
 1..501
 /organism="Zea mays"
 /cultivar="B73"
 /db_xref="taxon:4577"
 /clone="MEST8-G12"
 /clone_lib="ISUM3-TL"
 /tissue_type="Seedling and silk"
 /lab_host="DH10B"

/note="Vector: pT73PAC; Site_1: EcoRI; Site_2: NotI; ds-cDNA molecules were generated as follows. First-strand cDNA was prepared from oligo-dT selected mRNA by priming with a NotI oligo-dT primer (5' AACTGGGAAGATTCGCGCGCAGGATTTTTTTTTTTTTT). The resulting DNA-RNA hybrid was treated with RNase H and used as a template for DNA PolI-catalyzed second strand synthesis. After the addition of EcoRI adaptors, the ds-cDNAs were digested with NotI and size-selected. The resulting molecules were directionally cloned into the EcoRI and NotI sites of the pT73PAC vector."

BASE COUNT
 114 a 135 c 105 g 147 t

ORIGIN
 Query Match 3.1%; Score 36.8; DB 12; Length 501;
 Best Local Similarity 56.7%; Pred. No. 25;
 Matches 68; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

QY 384 TCTTCTGATGACATGAGGATCCGCCGCTCTCACTGGCGATAACGGGACGCGGAG 443
 Db 382 TCTTGGGTATATATCAGGTATCTCTCCACAGCAGGACGAGTGGGACTCCAGAA 441

QY 444 ACTGACGCTTCAGCGAGTACCATACCAACGCTTCTATTTCCGAAGCATGACCAACCA 503
 Db 442 GCCATGGACTCAGCAGCTACTTGCAGGAGGCTCGGATTCCTGAAGCATTCGCAATACA 501

RESULT 24
 BF586903/3
 LOCUS
 DEFINITION
 FM1_31_H07_g1_A003 Floral-Induced Meristem 1 (FM1) Sorghum propinquum cDNA, mRNA sequence.

ACCESSION
 BF586903
 VERSION
 BF586903.1 GI:11679227
 EST.
 SOURCE
 Sorghum propinquum.
 ORGANISM
 Sorghum propinquum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Sorghum.

REFERENCE
 1 (bases 1 to 540)
 Cordonnier-Pratt,M.-M., Gingle,A., Sudman,M., Marsala,C. and Pratt ,L.H.
 An EST database from Sorghum: floral-induced meristems Unpublished (2000)
 Contact: Cordonnier-Pratt MM
 Laboratory for Genomics and Bioinformatics
 The University of Georgia, Department of Plant Biology
 Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
 Tel: 706 542 1860
 Fax: 706 583 0210
 Email: mmpratt@uga.edu

Sequences have been trimmed to exclude PolyA, vector and regions below Phred quality 16. The threshold for highest quality sequence is 20.
 Seq primer: PolyTMix
 High quality sequence start: 9
 High quality sequence stop: 529
 POLYA-NO.

FEATURES
 Source
 Location/Qualifiers
 1..540
 /organism="Sorghum propinquum"
 /db_xref="taxon:132711"
 /clone_lib="Floral-Induced Meristem 1 (FM1)"
 /note="Organ: Floral-induced meristems; Vector: pBluescript II from Lambda Zap II; Site_1: XhoI; Site_2: EcoRI; mature plants were placed in a growth chamber for 15 days with 16 hr darkness and 8 hr light (flowering is induced by short-day conditions); 16 days after being returned to the greenhouse under natural long days during late April/early May, meristems were harvested. The library was made from poly-A RNA in the cloning vector lambda ZAP II. Clones to be sequenced were prepared by mass excision."

BASE COUNT
 151 a 114 c 137 g 138 t

ORIGIN
 Query Match 3.1%; Score 36.6; DB 12; Length 540;
 Best Local Similarity 56.1%; Pred. No. 28;
 Matches 69; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

QY 371 CGTCCCTGTACATTTCTTCTGATGACATCAGGATCCGCCGCTCTCACTGGCGATAC 430
 Db 124 CTTACCCCTCTGCTGCTTGGGTATATATCAGGTATCCCTCCAGCAGGACGACAAAC 65

QY 431 GGGCAGCGCGGAGACTGACGCTTCAGCCAGTACCATACCAACGCTTCAATTTCCGAAGG 490
 Db 64 TGGGACTCCAGAGCCATGACTCCAGCAGCTACTTGCCTGGGGTCTCAGACTCTGAAGG 5

QY 491 CAT 493
 Db 4 CAT 2

RESULT 25
 BI968888
 LOCUS
 DEFINITION
 GM830006B20A01 Gm-r1083 Glycine max cDNA clone Gm-r1083-1946 3', mRNA sequence.

ACCESSION
 BI968888
 VERSION
 BI968888.1 GI:16343293
 EST.
 KEYWORDS
 soybean.
 SOURCE
 Glycine max
 ORGANISM
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.

REFERENCE
 1 (bases 1 to 672)
 Vodkin,L., Keim,P., Shoemaker,R., Retzel,E., Khanna,A., Coryell,V., Erpelting,J., Rapp,C., Shoop,E., Pardinas,J., Liu,L. and Lewin,H.
 A Functional Genomics Program for Soybean (NSF 9872565)
 Unpublished (1999)
 Other ESTs: AW203245 corresponding to Gm-cl028-1235 (5')
 Contact: Vodkin, L.O., PI, A Functional Genomics Program for Soybean (NSF 9872565)
 Lewin, H. A., Director, Keck Center for Comparative and Functional Genomics
 University of Illinois
 Edwin R. Madigan Building, 1201 W. Gregory, Urbana, IL 61801, USA
 Tel: (217) 244-6147
 Fax: (217) 333-4582
 Email: l-vodkin@uiuc.edu

This clone is available through: Incyte Genomics, 4633 World Parkway Circle St. Louis, Missouri 63134. Phone (800) 430-0030 or (314) 427-3222 FAX: (314) 427-3324. Web site: <http://www.incyte.com/reagents/catalog.jsp?page=clones/collaboration/index>
 Seq primer: 5'-TTTTTTTTTTTTTTTTTTT(A/C/G)-3'.
 Location/Qualifiers
 1..672
 /organism="Glycine max"

```
/db_xref="taxon:3847"
/clone="Gm-r1083-1946"
/clone_lib="Gm-r1083"
```

/note--The library Gm-r1083 is a sequence-driven, rereacked set of 4,992 clones selected from cDNA libraries from various tissues and stages of development of soybean. It represents 1117 sequences from the progenitor library Gm-cl009 (from mature roots of 2 month old greenhouse grown 'Williams' soybean plants); 820 sequences from the progenitor library Gm-cl013 (from 2 to 3 week old whole plants of Williams); and 3055 sequences from library Gm-cl028 (from 'Supernod' plants whose seedlings were inoculated with *Bradyrhizobium japonicum*, courtesy of Dr. Gary Stacey). The 5' ESTs of the source clones from the different progenitor libraries was used to select singletons, or a representative of each contig, which were rereacked to form library Gm-r1083. The cDNA clones of the rereacked Gm-r1083 library were then sequenced at the 3' end. The contig analysis to select unique genes was performed by the laboratory of Ernest Retzel, Center for Computational Genomics and Bioinformatics, University of Minnesota, <http://web.ahc.umn.edu/biodata/nfssoy/>. . . Rereacking was performed by Incyte Genomics, St. Louis, <http://www.incyte.com>, and 3' sequencing by the Beck Center for Comparative and Functional Genomics, University of Illinois, <http://www.lie.uic.edu/biotech/beck.html>. Note: The corresponding 5' EST from each clone in the Gm-r1083 library is listed in the "OTHER EST" field. The detailed information on the source library for each clone can also be obtained by referring to the Incyte Genomics clone ID of the original cDNA library that is also listed under "OTHER EST".

BASE COUNT	167 a	166 c	140 g	174 t	25 others
ORIGIN					

	Query Match	3.1%	Score 36.4;	DB 13;	Length 672;
	Best Local Similarity	53.3%;	Pred. NO. 33;		
	Matches	73;	Conservative 0;	Mismatches 64;	Indels 0; Gaps 0;
Qy	391	GAATGACATCAGGATCCCGCCGTC	CTCACTGCGGATACGGGCGACGCCGGAGACATGACG	450	
Db	465	GGATTATATCAGGAATACCTCCAGCANNCCGGCCACCAAGGTATCCCTTGAAGACATCG	524		
Qy	451	CTTCAGCCAGTACCATCAACCAAGCGTTCA	TTTCCGAAGGCATGACCACACACATCTGSCAA	510	
Db	525	CCTCCAAACTACAAACCAGCGGCTCTGAC	TCTGNAAGCGATGACAAATACATCTCCAC	584	
Qy	511	TCCGGTAGACCGGTAAC	527		
Db	585	TGGCATATGCTTGGGAC	601		

RESULT 26	BI498039/c	BI498039	499 bp	mRNA	linear	EST 29-NOV-2001
LOCUS		sag1sc08.y1	Gm-c1080	Glycine max	cDNA clone	GENOME SYSTEMS CLONE
DEFINITION		ID: Gm-c1080-1263	5'	similar to	TR:Q9R6U1 Q9R6U1	SQDX. ; , mRNA sequence.

ACCESSION	BI498039	sequence.
VERSION	BI498039.1	GT:15337383

RETICULOS	SOURCE	ORGANISM
251.	soybean.	<i>Glycine max</i>
		Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; <i>Glycine</i> .

REFERENCE

AUTHORS

1 (bases 1 to 499)

Slycine.

Shoenaker, R., Reim, P., Vodka, L., Erpelting, J., Corvelli, V., Khanna
A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C.,
Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers,
Y., Person, B., Swallier, T., Gibbons, M., Pape, D., Harvey, N., Schurk,
R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann

, R., Waterston, R. and Willson, R.
Public Soybean EST Project
Unpublished (1999)
Contact: Shoemaker R/Public Soy
Public Soybean EST Project
Washington University School of
4444 Forest Park Parkway, Box 8
Tel: 314 286 1800

This clone is available through: Resgen, Invitrogen Corp. 2130 South Memorial Parkway Huntsville, AL 35801 For further information call: (800)-533-4363 or contact via email: ccu@resgen.com High quality sequence stop: 327.

FEATURES

source

```

/organism="Glycine max"
/db_xref="taxon:3847"
/clones="GENOME SYSTEMS CLONE ID: Gm-c1080-1263"
/clone.lib="Gm-c1080"
/tissue_type="Roots of 8 day old 'bragg' supernatant mutant NTS382 seedlings"
/dev_stage="8 days old"
/lab_host="DH10B"

```

/note=Vector: pBluescript II SK+, Site_1: EcoRI; Site_2: XhoI; The mRNA was isolated from roots of 8 day old 'Bragg' supernodulating mutant NRS382 seedlings that were infected with Bradyrhizobium japonicum, strain USDA 110, 72 hours prior to harvest. Dr. Gary Stacey generously donated the tissue. The roots were flash-frozen in liquid nitrogen. StrataGene's cDNA Synthesis Kit (catalog number 200401) was used to synthesize the cDNA. First-strand synthesis was performed with 5-methyl dCTP, hence the ligated cDNA was hemimethylated. A modification of StrataGene's first-strand synthesis primer was used. An 'anchor' nucleotide (V=A, C, or G) was added to the 3' end of the primer (GAGAGAGAGAGAGAGACTACTGCGAG(7)18V) to anchor the primer at the 5' end of the poly(A) tract. After second-strand synthesis, the cDNA ends were filled in with cloned pfu DNA, ligated to EcoRI adapters and subsequently phosphorylated. The cDNA was then precipitated and redissolved in sterile, RNase-, DNase-free water. The XhoI site within the first- strand synthesis primer was then restricted by digestion with XhoI from Promega (40U/ul); all XhoI sites in the cDNA would be protected by their hemimethylated status. The cDNA constructs were size-fractionated with a 500bp cutoff, using Sephacryl S-500 High Resolution (Pharmacia Biotech) in a 2-mm diameter column and a bed volume of approximately 1ml. The column eluent was precipitated, redissolved, and ligated into StrataGene's pBluescript II XR predigested vector (pBluescript II SK(+) vector that has been digested with EcoRI and XhoI, and phosphorylated by StrataGene). This library was constructed in the laboratory of Dr. Paul Keim and Dr. Virginia H. Coryell at Northern Arizona University.

89 c 137 g 130 t.

BASE COUNT
ORIGIN

Query Match	3.1%	Score 36.2;	DB 13;	Length 499;
Best Local Similarity	54.0%;	Pred. No. 36;		
Matches 74;	Conservative	0;	Mismatches	63;
			Indels	0;
			Gaps	0;

Qy	391	GAATGACATCAGGGATCCGCCGCTCTCACTGGCGATAACGGGCACGCCGGAGACTGACG	450
Db	368	GGATTATATCAGGAATACCTCCACGACCTGCCGCCACACAGGATATCCTCTGAACATCC	309

[illegible]

Db 513 CATTGCAACAC 524

RESULT 29
CNS031JD
LOCUS
DEFINITION
Tetraodon nigroviridis genome survey sequence PUC-ori end of clone 029c12 of library G from Tetraodon nigroviridis, genomic survey sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
CNS031JD 849 bp DNA linear GSS 17-MAY-2000
Tetraodon nigroviridis genome survey sequence
AL245650.1 GI:7966662
GSS: genome survey sequence.
Tetraodon nigroviridis
Tetraodon nigroviridis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes; Tetraodontidae; Tetraodon.
1 (bases 1 to 849)
Roest-Crolius, H., Jaillon, O., Dasilva, C., Bouneau, L., Fisher, C., Bernot, A., Fizames, C., Wincker, P., Brottier, P., Quetier, F., Saurin, W. and Weissenbach, J.
Human gene number estimate provided by genome wide analysis using Tetraodon nigroviridis DNA sequence
Unpublished
2 (bases 1 to 849)
Roest-Crolius, H., Jaillon, O., Dasilva, C., Fizames, C., Fisher, C., Bouneau, L., Billault, A., Quetier, F., Saurin, W., Bernot, A. and Weissenbach, J.
Characterization and repeat analysis of the compact genome of the freshwater pufferfish Tetraodon nigroviridis
Unpublished
3 (bases 1 to 849)
Genoscope.
Direct Submission
Submitted (12-APR-2000)
This sequence is a single read and was generated as part of a large scale clone-end and sequencing project of the Tetraodon nigroviridis genome. For more information, please take a look at
http://www.genoscope.cns.fr/Tetraodon.
Location/Qualifiers
1. .849
/organism="Tetraodon nigroviridis"
/db_xref="taxon:99883"
/clone_lib="G"
/clone="029c12"
/note="Genoscope sequence ID : C0BG029BB06SP1-end : PUC-ori"
PUC-ori
BASE COUNT 176 a 228 c 274 g 159 t 12 others
ORIGIN
Query Match 3.0%; Score 35.8; DB 17; Length 849;
Best Local Similarity 50.9%; Pred. No. 50;
Matches 85; Conservative 0; Mismatches 82; Indels 0; Gaps 0;
QY 518 GACCGTAGCTGGGAAAGGCGACCTGCCATTACACATCTCCGCTCATTCGCCAGTG 577
DB 231 GGCCTGGGACGGGGGACAAAGGACTGCGCCACCTGCAGGAGGCGGCTCTCAACGAC 290
QY 578 TTCTGTCTGTGACGACGACGTGCTTGTATTTCTACGCGCGGCGCCACACGAGCCA 637
DB 291 CTGCACTGCTGACACGGCTGCGTCTGCTGCCAGAACACAGGCGCGCCACCCACCC 350
QY 638 GCGAATGATTTCCTTCATCTTCAGCTGATACATACACACGACCA 684
DB 351 GCGCGCGCTAGCTGCGGTAGCTGCTGTGTACGATAAAGAACCA 397

RESULT 30
BQ878757/c
LOCUS
DEFINITION
AGENCOURT_8121988 Lupski_dorsal_root_ganglion Homo sapiens cdna clone IMAGE:6179307 5', mRNA sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
BQ878757 1 GI:22270765
EST.
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 974)
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs@mail.nih.gov
Tissue Procurement: Dr. James R. Lupski
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM13560 row: e column: 04
High quality sequence start: 62
High quality sequence stop: 560.
Location/Qualifiers
1. .974
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6179307"
/clone_lib="Lupski_dorsal_root_ganglion"
/sex="male"
/tissue_type="dorsal root ganglia"
/dev_stage="adult, 36 yr"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6 (Life Technologies); Site_1: NotI; Site_2: SalI; cDNA made by oligo-dT priming. Directionally cloned using the following adaptors: 5'-TCGACCCAGCGCTCG-3' and 5'-GACTGATTCTAGTCGAGCGCGCCCTT(15)-3'. Size selected > 1 kb for average insert length 1.7 kb. This is a primary library, non-amplified. Library constructed by Life Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor College of Medicine) and is available through Life Technologies."
228 a 246 c 256 g 244 t
BASE COUNT
ORIGIN
Query Match 3.0%; Score 35.8; DB 14; Length 974;
Best Local Similarity 65.8%; Pred. No. 51;
Matches 52; Conservative 0; Mismatches 27; Indels 0; Gaps 0;
QY 525 AACGCTGGGAAAGGCGACCTGCCATTACACATCTCCGCTCATTCGCCAGGTGTGTGTC 584
DB 478 AATTCCTAGACCAAGCAGCAGCTTCCAGGAACAGAGCTCCGCCACTCACCTGGTGACCTTC 419
QY 585 TCGTACCGCAGACGTGCTT 603
DB 418 CTCGGGGAAGACTGCTT 400

RESULT 31
CNS017IM/c
LOCUS
DEFINITION
Drosophila melanogaster genome survey sequence T7 end of BAC BACN17E01 of DrosBAC library from Drosophila melanogaster (fruit fly), genomic survey sequence.
AL108088
1 (bases 1 to 948)
GSS.
AL108088.1 GI:5628392
Drosophila melanogaster.
Drosophila melanogaster.
Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

4

kb (range 1.0-4.0 kb). 15/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA). Note: this is a NIH_MGC Library."

```

BASE COUNT      272 a 166 c 184 g 209 t
ORIGIN

Query Match      3.0%; Score 35.4; DB 12; Length 831;
Best Local Similarity 52.3%; Pred. No. 65;
Matches 78; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

QY 1004 TCCTGCTATATTAATGTCAGCGCCACTATGCGCAGATGACGAGATCAGGTTTAAATCCCC 1063
DB 386 TCTGTTTCATCTGTCAGATGCGCATCTCAAGATCTCTCGACAGATCCCACTCCCA 327
QY 1064 GATAATCCGTCGAGTCTGAGGATGGAAGGCTGCTTCTCTGAAAGGAATAAA 1123
DB 326 GAGGATTGCTTAAACAGGAGTCAVAAGCTTGATGAGTTATGATAAAAGATGTTTCATC 267
QY 1124 AGTCACATCATGCCCTCTTTTCTGCGTT 1152
DB 266 ATGGAATTTATGTCCTTTTGTTCACTT 238

RESULT 34
AV987761
LOCUS
DEFINITION
AV987761 Nori Satoh unpublished cDNA library, tailbud embryo Ciona
Intestinalis cDNA clone citb18121 5', mRNA sequence.
ACCESSION
AV987761
VERSION
AV987761
KEYWORDS
EST.
SOURCE
Ciona intestinalis.
ORGANISM
Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
Phlebobranchia; Clonidae; Ciona.
REFERENCE
1 (bases 1 to 582)
Satoh,N., Satou,Y., Kohara,Y. and Shin-I,T.
Expressed genes in Ciona intestinalis
Unpublished (2000)
JOURNAL
Contact: Nori Satoh
Department of Zoology
Kyoto University
Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
Tel: 81-75-753-4081
Fax: 81-75-705-1113
Email: satoheascidian.zool.kyoto-u.ac.jp.
FEATURES
Location/Qualifiers
1..582
/organism="Ciona intestinalis"
/db_xref="taxon:7719"
/clone="citb18121"
/clone_lib="Nori Satoh unpublished cDNA library, tailbud
embryo"
/tissue_type="whole animal"
/dev_stage="tailbud embryo"
/note="vector: pBluescript SK"
BASE COUNT      140 a 136 c 133 g 173 t
ORIGIN

Query Match      3.0%; Score 35.2; DB 10; Length 582;
Best Local Similarity 53.7%; Pred. No. 71;
Matches 73; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

QY 838 TTACCGGTGTCCTGACACCTTCGCCCATCATGATCGCGCATCGGTGAGATAGGCACAA 897
DB 435 TTACTGTTGATGAGTCTGACTCACTAAGACCCCAACAGCATCAACGGTTGAGATGGGTAAA 494
QY 898 CAATGAATCACACAGATAATTCAGGGAACAGCTTCCTGCTTACGGGTGATGAGTTT 957
DB 495 TAATGTTTCCCCCAAGGAGTGATGGCAAAATGCTCTTATACCTATTTTITAGGCTT 554
QY 958 TTTGTCGACAAATAGT 973

```

```

DB 555 ATTGTGAGATGCTGGT 570
||||| || |||
RESULT 35
CNS00LT2/c
LOCUS
DEFINITION
CNS00LT2 1101 bp DNA linear GSS 14-JUN-1999
Drosophila melanogaster genome survey sequence TET3 end of BAC:
BACR48P19 of RPCI-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION
AL078714.1 GI:5102004
VERSION
AL078714.1
KEYWORDS
GSS.
SOURCE
Drosophila melanogaster.
ORGANISM
Drosophila melanogaster.
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE
1 (bases 1 to 1101)
Genoscope.
AUTHORS
Direct Submission
TITLE
Submitted (11-JUN-1999) Genoscope - Centre National de Sequencage :
JOURNAL
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT
Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see http://www.fruitfly.org The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Osoegawa and
Aaron Mamoser in Pieter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCI-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain y2; cn bw sp, the same strain used for the BDGP's
pi and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES
Location/Qualifiers
1..1101
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone="BACR48P19"
/clone_lib="RPCI-98"
/note="end : TET3"
BASE COUNT      469 a 6 c 69 g 151 t 406 others
ORIGIN

Query Match      3.0%; Score 35.2; DB 17; Length 1101;
Best Local Similarity 15.4%; Pred. No. 76;
Matches 47; Conservative 125; Mismatches 134; Indels 0; Gaps 0;

QY 434 CAGCGCGGAGACTGACGCTTCAGCCAGTACCATACCAAGCTTCATTTCCGAGGCAT 493
DB 1055 CMCCMCMCHMCAAMMMHMTTHMMMMMMAMMMAMMMAMMMAMMMAMMMAMMM 996
QY 494 GACACACACACTGGCAATCCGTAAGCGGTGAGAAAGGCGCATCAATAA 553
DB 995 MHHMMMMMMATTHAWHTTTTTHMMAMCMTHMMMMMMAMMMAMMMAMMMAMMM 936
QY 554 CACATCTCCGCTCATTCGCCAGGTGTTCTGTCTGCTGACGAGAGCTTCGTATTTCT 613
DB 935 CMMHMMCCCMCMAMMMAMMMMTTTHHMMCCMCMCMCMCMCMAMMMAMMMAMMTT 876
QY 614 ACSCCGGCGGCCACACGAGCGAGCAATGATTTCCCTTCATCTTCAGCTGATACAA 673
DB 875 THMCCMCMMMMMAMMMAMMMAMMMAMMMAMMMAMMMAMMMAMMMAMMMAMMTT 816
QY 674 TACACGCGCATAAATTCATGTCCTTTTTCGGGAGCTAGCATCCCACTGAACGATAAG 733
DB 815 MHHMMCMAMMAAAATTTMMMMAMMMAMMMAMMMAMMMAMMMAMMMAMMMAMMM 756

```


LOCUS BO6111687 445 bp mRNA linear EST 26-JUN-2002
DEFINITION Gap3901.y1 Gm-cl087 Glycine max cDNA clone SOYBEAN CLONE ID:
Gm-cl087-5306 5' similar to SW:SUSY_SOYBN P13708 SUCROSE SYNTHASE
, mRNA sequence.
ACCESSION BO6111687
VERSION 1
KEYWORDS GI:21601356
SOURCE EST
ORGANISM soybean.
Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.
REFERENCE 1 (bases 1 to 445)
AUTHORS Shoemaker,R., Keim,P., Vodkin,L., Erpelding,J., Coryell,V., Khanna
A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,
Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers
Y., Person,B., Swallier,T., Gibbons,M., Pape,D., Harvey,N., Schuck
R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann
R., Waterston,R. and Wilson,R.
TITLE Public Soybean EST Project
JOURNAL Unpublished (1999)
COMMENT Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available through: ResGen, Invitrogen Corp. 2130
South Memorial Parkway Huntsville, AL 35801 For further information
call: (800)-533-4363 or contact: ccu@resgen.com web site:
www.resgen.com
Seq primer: -40RP from Gibco.

FEATURES

```

1. 445
/organism="Glycine max"
/db_xref="taxon:3847"
/clone="SOYBEAN CLONE ID: Gm-cl087-5306"
/clone_lib="Gm-cl087"
/tissue_type="Roots"
/lab_host="DH10B"
/notes="Vector: pBluescript II SK+; Site1: EcoRI; Site2: XhoI; The mRNA was prepared using poly(I)tract mRNA system from PROMEGA. The cDNA was prepared using the STRATAGENE kit. Complementary DNA was synthesized from mRNA using a primer consisting of a poly(dT) sequence with a XhoI restriction site
(5'GAGAGAGAGAGAGAGAGAACTAGTCGAGCTTTTTCCTTTTTTTTTT). EcoRI adapters (5'OH-AATTGGCCACGAG and 3'GCCGTCGCTp) were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction sites of the pBluescript vector. The vector was previously dephosphorylated. The ligated cDNA fragments were transformed into DH10B by electroporation. Library is in LB medium with ampicillin and glycerol 8%. Average insert size: 800 bp."
100 c 135 g 105 t

```

[illegible]

	Query Match	2.9%	Score 34.8;	DB 14;	Length 445;
	Best Local Similarity	55.9%	Pred. No. 89;		
	Matches	66;	Conservative	0;	Mismatches 52; Indels 0; Gaps 0;
QY	387	TTCTGAATGACATCAGGATCCGCCGCTCTCACTGGCGATACGGCGACGCCGGAGACT	446		
Db	338	TGCACAAATGATCTCAGCAGGACCAACCATTCGATGTGGCGAATGTGGCAAGCCGCAAGTC	279		
QY	447	GAGGCTTCACCCAGTACCATACCAAGCGCTTCATTTCCGAAGGCATGACCAACCAAC	504		
Db	278	ATGCGCTCAACCATGTCAACCAAGAAAGCCTCGTATACAGCAGGCTGCACGAAGAC	221		

RESULT 39	AW459606	446 bp	linear	EST 03-DEC-2001
LOCUS	sh9c12.v1	Gm-cl016	Glycine max	CDNA clone
DEFINITION	Gm-cl016-7343.v1	similar to	TR:024302	O24302 SUCROSE SYNTHASE ; , mRNA sequence.
ACCESSION	AW459606			
VERSION	AW459606.1	GI:7029823		
KEYWORDS	EST.			
SOURCE	soybean.			
ORGANISM	Glycine max			
REFERENCE	1. (bases 1 to 446)			
AUTHORS	Shoemaker, R., Keim, P., Vodkin, L., Erpelding, J., Corvelli, V., Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schur, R., Ritter, E., Kohn, S., Shih, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.			
TITLE	Public Soybean EST Project			
JOURNAL	Unpublished (1999)			
COMMENT	Contact: Shoemaker R/Public Soybean EST Project Public Soybean EST Project Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA Tel: 314 286 1800 Fax: 314 286 1810 Email: est@watson.wustl.edu This clone is available through: ResGen, Invitrogen Corp. 2130 South Memorial Parkway Huntsville, AL 35801 For further information call: (800)-533-4363 or contact via email: ccu@resgen.com Insert length: 1002 Std Error: 0.00 Seq primer: -40RP from Gibco High quality sequence stop: 409.			

FEATURES
SOURCE

1. 446
/organism="Glycine max"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-cl016-7343"
/clone_lib="Gm-cl016"
/tissue_types="Immature flowers of field grown plants"
/lab_host="Xl10-Gold"
/note="Vector: pBluescript II XR; Site_1: EcoRI; Site_2: XhoI; This cDNA library was constructed from mRNA isolated from Immature flowers of field grown plants. The cDNA library was prepared using the Stratagene pluescript II XR library construction kit. Complementary DNA was synthesized from mRNA using a primer consisting of a poly (dT) sequence with a XhoI restriction site. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pBluescript vector. The ligated cDNA fragments were transformed into Xl10-Gold host cells. This library was constructed by Dr. Randy Shoemaker and Dr. John Enold."/>

[illegible]

	Query Match	2.9%	Score 34.8:	DB 10;	Length 446;
	Best Local Similarity	55.9%;	Pred. No. 89;		
	Matches 66;	Conservative 0;	Mismatches 52;	Indels 0;	Gaps 0;
Qy	387	TTCTGAATGACATCAGGGATCCGCCCGTCTCACTGGCGATAACGGGCACGCCGGAGACT	446		
Db	334	TGCACAATGATCTCAGCAGGACCACCATTTCATGTGCGAATGTGTGCAAGCCGCRAGTC	275		
Qy	447	GAGCCTTCAGCCAGTACCATTACCAAACGCTTCATTTTCCGAAGCATGACCCACACAC	504		
Db	274	ATGGCCTCAACCATGTGTAACCAAAAAGCCTCGTGATACAGCAGGCTGCACGAAGCAC	217		

synthesize the cDNA. First-strand synthesis was performed with 5-methyl dCTP, hence the ligated cDNA was hemimethylated. A modification of Stratagene's first-strand synthesis primer was used. An 'anchor' nucleotide (V=A, C, or G) was added to the 3' end of the primer [GAGAGAGAGAGAGAGACTAGTCTCGAG(T)18V] to anchor the primer at the 5' end of the poly(A) tract. After second strand synthesis, the cDNA ends were filled in with cloned Pfu DNA, ligated to EcoRI adapters and subsequently phosphorylated. The cDNA was then precipitated and redissolved in sterile, RNase-, DNase-free water. The XhoI site within the first-strand synthesis primer was then restricted by digestion with XhoI from Promega (40U/ul); all XhoI sites in the cDNA would be protected by their hemimethylated status. The cDNA constructs were size-fractionated with a 500 bp cutoff, using GibcoBRL Life Technologies' cDNA Size Fractionation column. The column eluent was then precipitated, redissolved, and ligated into Stratagene's pluscript II XR Predigested vector (pBluescript II SK(+)) vector that has been digested with EcoRI and XhoI, and phosphorylated by Stratagene). 100% of the white and blue colonies appear to contain recombinant plasmids with cDNA inserts, based on size (n=24 and 6, respectively). This library was constructed by Dr. Paul Keim and Dr. Virginia Coryell.

```

BASE COUNT      106 a   130 c   134 g   137 t
ORIGIN
Query Match      2.9%; Score 34.8; DB 10; Length 507;
Best Local Similarity 55.9%; Pred. No. 91;
Matches 66; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

QY 387 TTCTGAATGACATFACGGATCCGCCGCTCTCACTGGCGATACGGCGACGGCGAGACT 446
Db 199 TGCACAATGATCTCAGCAGGACCAACCATTCGATGTGGCAATGTTGGCAAGCGCAAGTC 140
QY 447 GACGCTTCAGCCAGTACCATACCAACGCTTCAATTTCCGAGGCGATGACCACACAC 504
Db 139 ATGGCCTCAACCACTGTCAACCAAAAGCGCTGTATACAGCAGGCTGCACGAAGCAC 82

RESULT 42
A0785148
LOCUS
DEFINITION
HS_3159_AL_B06_T7C CIT Approved Human Genomic Sperm Library D Homo
sapiens genomic clone Plate-3159 Col-11 Row=C, DNA sequence.
ACCESSION
A0785148
VERSION
A0785148.1 GI:5692772
KEYWORDS
GSS.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 512)
AUTHORS
Mahairas, G.G., Wallace, J.C., Smith, K., Swartzell, S., Holzman, T.,
Keller, A., Shaker, R., Furlong, J., Young, J., Zhao, S., Adams, M.D. and
Hood, L.
TITLE
Sequence-tagged connectors: A sequence approach to mapping and
scanning the human genome
Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
JOURNAL
99380589
MEDLINE
High Throughput Sequencing Center
COMMENT
Contact: Mahairas GG, Wallace JC, Hood L
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Clones are derived from the human BAC library RPCI-11. For BAC
library availability, please contact Pieter de Jong
(pieterdejong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm)
or from Research Genetics (info@resgen.com). BAC end Web Server:
http://www.htsc.washington.edu
Plate: 3159 row: C column: 11
Seq primer: T7

A0785148
LOCUS
DEFINITION
HS_3159_AL_B06_T7C CIT Approved Human Genomic Sperm Library D Homo
sapiens genomic clone Plate-3159 Col-11 Row=C, DNA sequence.
ACCESSION
A0785148
VERSION
A0785148.1 GI:5692772
KEYWORDS
GSS.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 512)
AUTHORS
Mahairas, G.G., Wallace, J.C., Smith, K., Swartzell, S., Holzman, T.,
Keller, A., Shaker, R., Furlong, J., Young, J., Zhao, S., Adams, M.D. and
Hood, L.
TITLE
Sequence-tagged connectors: A sequence approach to mapping and
scanning the human genome
Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
JOURNAL
99380589
MEDLINE
High Throughput Sequencing Center
COMMENT
Contact: Mahairas GG, Wallace JC, Hood L
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Clones are derived from the human BAC library RPCI-11. For BAC
library availability, please contact Pieter de Jong
(pieterdejong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm)
or from Research Genetics (info@resgen.com). BAC end Web Server:
http://www.htsc.washington.edu
Plate: 3159 row: C column: 11
Seq primer: T7

```

```

Class: BAC ends
High quality sequence stop: 512.
Location/Qualifiers
1..512
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="Plate-3159 Col-11 Row=C"
/clone_lib="CIT Approved Human Genomic Sperm Library D"
/sex="male"
/note="Organ: sperm; Vector: pBelOBAC11; BAC Clones in
E-Coli DH10B"
BASE COUNT      127 a   124 c   117 g   133 t   11 others
ORIGIN
Query Match      2.9%; Score 34.8; DB 17; Length 512;
Best Local Similarity 52.0%; Pred. No. 91;
Matches 78; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

QY 900 ATGAATACACAGATTAATTCAGGGAACAGTTCTGCTTACGGGTGATGAGGTTTTT 959
Db 101 ATCCGAGCACACAGCTGGCTAAGTGGCAGACTATTGTCTACTGTTATTTAGGTTTC 160
QY 960 TGTCTGCATAATAGTGAAGCGGTGACAGCATATCAGACGGCTCAGTCTGCTATATTACTG 1019
Db 161 TGTGTGCACCGAGGTGGCTGGGTTATATAATGCCACCTGACCCCAAGATCTGTTCTCAG 220
QY 1020 TCATGCCCACTATGGCAGATGACCAATCA 1049
Db 221 ATGTGCACCTGTGTGGCTGCCAAGACCA 250

RESULT 43
A0662456
LOCUS
DEFINITION
HS_5455_A2_E02_T7A RPCI-11 Human Male BAC Library Homo sapiens
genomic clone Plate-1031 Col-4 Row=I, DNA sequence.
ACCESSION
A0662456
VERSION
A0662456.1 GI:5170224
KEYWORDS
GSS.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 514)
AUTHORS
Mahairas, G.G., Wallace, J.C., Smith, K., Swartzell, S., Holzman, T.,
Keller, A., Shaker, R., Furlong, J., Young, J., Zhao, S., Adams, M.D. and
Hood, L.
TITLE
Sequence-tagged connectors: A sequence approach to mapping and
scanning the human genome
Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
JOURNAL
99380589
MEDLINE
High Throughput Sequencing Center
COMMENT
Contact: Mahairas GG, Wallace JC, Hood L
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Clones are derived from the human BAC library RPCI-11. For BAC
library availability, please contact Pieter de Jong
(pieterdejong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm)
or from Research Genetics (info@resgen.com). BAC end Web Server:
http://www.htsc.washington.edu
Plate: 1031 row: I column: 4
Seq primer: T7
Class: BAC ends
High quality sequence stop: 514.
Location/Qualifiers
1..514
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="Plate-1031 Col-4 Row=I"

```

/clone_lib="RPCI-11 Human Male BAC Library"

/sex="male"

/note="Vector: pBACE3.6; Site_1: EcoRI; Site_2: EcoRI;
Male blood DNA was isolated from one randomly chosen donor
and partially digested with a combination of EcoRI and
EcoRI Methylase. Size selected DNA was cloned into the
pBACE3.6 vector at EcoRI sites"

BASE COUNT 123 a 130 c 125 g 132 t 4 others

Query Match 2.9%; Score 34.8; DB 17; Length 514;

Best Local Similarity 52.0%; Pred. No. 91;

Matches 78; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

QY 900 ATGAATACACAGATAATTCAGGAAACGTTCTGCTTACGGGTGATGAGTGTTC 959

DB 82 ATCCGACACACAGCTGGCTAAAGTGGCAGACTTATTCTTACTGTTATTTAGGTTC 141

QY 960 TGTCTGACATAGTGAAGCGGTGACAGCATATCAGAGCGTCTAGTCTGCTGATATTACTG 1019

DB 142 TGTGTGCACCGAGTGGCTGGGTATTAATGCAGTGCACCAAGATACTGTTCTTCAG 201

QY 1020 TCATGGCCACTATGGCAGATGACAGATCA 1049

DB 202 ATGTGCACCGCTGTGTGGCTGCCAAGACCA 231

RESULT 44

BE440931/c

LOCUS

DEFINITION BE440931 523 bp mRNA linear EST 04-DEC-2001

sp49g12.y1 Gm-cl043 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:

Gm-cl043-1631 5' similar to SW:SUSY_SOYBN P13708 SUCROSE SYNTHASE

; mRNA sequence.

ACCESSION BE440931

VERSION BE440931.1 GI:9440426

KEYWORDS EST.

SOURCE soybean.

ORGANISM Glycine max

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;

Glycine.

1 (bases 1 to 523)

Shoemaker,R., Keim,P., Vodkin,L., Erpellding,J., Corvelli,V., Khanna

,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,

Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers

,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk

,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann

,R., Waterston,R. and Willson,R.

Public Soybean EST Project

Unpublished (1999)

Contact: Shoemaker R/Public Soybean EST Project

Public Soybean EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

This clone is available through: ResGen, Invitrogen Corp. 2130

South Memorial Parkway Huntville, AL 35801 For further information

call: (800)-533-4363 or contact via email: ccu@resgen.com

High quality sequence stop: 485.

Location/Qualifiers

1. 523

/organism="Glycine max"

/db_xref="taxon:3847"

/clone_lib="Gm-cl043"

/tissue_type="Hypocotyl and Plumule, germinating seeds"

/lab_host="DH10B"

/note="Vector: pT7T3Pac (Pharmacia); Site_1: EcoRI;

Site_2: NotI; This cDNA library was constructed from mRNA

isolated from hypocotyl and plumule tissues of seeds

FEATURES

SOURCE

1. 523

/organism="Glycine max"

/db_xref="taxon:3847"

/clone_lib="Gm-cl043"

/tissue_type="Hypocotyl and Plumule, germinating seeds"

/lab_host="DH10B"

/note="Vector: pT7T3Pac (Pharmacia); Site_1: EcoRI;

Site_2: NotI; This cDNA library was constructed from mRNA

isolated from hypocotyl and plumule tissues of seeds

germinated for three days of the cultivar Williams.

Complementary DNA was synthesized from mRNA using a primer

consisting of a poly(dT) sequence with a NotI restriction

site. EcoRI adapters were ligated to the blunt-ended cDNA

fragments followed by digestion with EcoRI and NotI. The

cDNA fragments were directionally cloned into the

EcoRI-NotI restriction site of the pT7T3-pac vector. The

ligated cDNA fragments were transformed into DH10B host

cells (Gibco BRL). This library was constructed by Dr.

Randy Shoemaker.

BASE COUNT 115 a 132 c 139 g 137 t

Query Match 2.9%; Score 34.8; DB 10; Length 523;

Best Local Similarity 55.9%; Pred. No. 91;

Matches 66; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

QY 387 TTCTGAATGACATCATGAGGATCCGCCCTCTCTCATCTGGCGGTAACGGGCACCGCGAGACT 446

DB 228 TGCACAATGATCTCAGCAGGACCACTTGTGGGGAATGTTGGCAAGCCGCAAGTC 169

QY 447 GAGCGTTCAGCCAGTACCATACCAACGCTTCATTTCCGAAGGCATGACCAACACAC 504

DB 168 ATGGCTCAACCACTGTCAACCAAGCCCTGTAACAGGCTGTAACAGGCTGCAACCAACAC 111

RESULT 45

AW279073/c

LOCUS

DEFINITION AW279073 527 bp mRNA linear EST 02-DEC-2001

sg07d10.y1 Gm-cl019 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:

Gm-cl019-4940 5' similar to SW:SUSI_SOYBN P13708 SUCROSE SYNTHASE

; mRNA sequence.

ACCESSION AW279073

VERSION AW279073.1 GI:6667622

KEYWORDS EST.

SOURCE soybean.

ORGANISM Glycine max

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;

Glycine.

1 (bases 1 to 527)

Shoemaker,R., Keim,P., Vodkin,L., Erpellding,J., Corvelli,V., Khanna

,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,

Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers

,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk

,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann

,R., Waterston,R. and Willson,R.

Public Soybean EST Project

Unpublished (1999)

Contact: Shoemaker R/Public Soybean EST Project

Public Soybean EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

This clone is available through: ResGen, Invitrogen Corp. 2130

South Memorial Parkway Huntville, AL 35801 For further information

call: (800)-533-4363 or contact via email: ccu@resgen.com

Insert length: 329 Std Error: 0.00

Seq primer: -40RP from Gibco

High quality sequence stop: 433.

Location/Qualifiers

1. 527

/organism="Glycine max"

/db_xref="taxon:3847"

/clone_lib="Gm-cl019"

/tissue_type="Immature seed coats of greenhouse grown

plants"

/note="Vector: pSPORT1 (Life Technologies); Site_1: Not I;

FEATURES
source
1. .542
/organism="Glycine max"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-cl049-4322"
/clone_lib="Gm-cl049"
/tissue_type="whole seedlings of greenhouse grown plants"
Location/Qualifiers
High quality sequence stop: 421.
call: (800)-533-4363 or contact via email: ccu@resgen.com
southern MEMPHIS, 1474 W. HUNTSVILLE, AL 35801. for further information

```

/dev_stage="3 week old"
/lab_host="DH10B"
/note="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2:
XhoI; The Clark NIL was constructed and seed was provided
by Dr. J. Specht, University of Nebraska (Shoemaker and
Specht, 1995). The cDNA library was constructed from mRNA
isolated from whole seedlings of 3 week old greenhouse
grown plants. Complementary DNA was synthesized from mRNA
using a primer consisting of a poly(dT) sequence with a
XhoI restriction site and a 3' anchor. EcoRI adapters were
ligated to the blunt-ended cDNA fragments followed by XhoI
digestion. The cDNA fragments were directionally cloned
into the EcoRI-XhoI restriction site of the pBluescript
vector. The ligated cDNA fragments were transformed into
DH10B host cells (GibcoBRL). The library was constructed
in cooperation with Dr. Paul Keim's laboratory at Northern
Arizona University."
BASE COUNT      118 a      128 c      139 g      157 t
ORIGIN

```

```

Query Match      2.9%; Score 34.8; DB 13; Length 542;
Best Local Similarity 55.9%; Pred. No. 91;
Matches 66; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

QY 387 TTCTGAATGACATCAGGATCCGCCGCTCTACTGCGATACACGGCAGCGGAGACT 446
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 137 TGCACATGATCTCAGCAGGACCACTTCATGTCGTGGCGAATGTTGGCAAGCGCAAGTC 78
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 447 GAGCCTTCAGCAGTACCAACCAACGCTTCATTTCCGAAGGATGACCCACAC 504
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 77 ATGGCCTCAACCACTGTCAAAACCAAAAGCCTGTATACAGCAGGTCGCAAGAACAC 20
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

RESULT 48
BI427022/c
LOCUS
DEFINITION
sag10403.y1 Gm-cl080 Glycine max cDNA clone GENOME SYSTEMS CLONE
ID: Gm-cl080-558 5' similar to SW:SUSY_SOYBN P13708 SUCROSE
SYNTHASE ;, mRNA sequence.
BI427022
BI427022.1 GI:15204254
EST.
soybean.

```

```

ORGANISM
Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.
1 (bases 1 to 566)
Shoemaker,R., Keim,P., Vodkin,L., Erpelting,J., Coryell,V., Khanna
,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,
Wyllie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers
,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk
,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann
,R., Waterston,R. and Wilson,R.
Public Soybean EST Project
Unpublished (1999)
Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available through: Resgen, Invitrogen Corp. 2130
South Memorial Parkway Huntsville, AL 35801 For further information
call: (800)-533-4363 or contact via email: ccu@resgen.com
High quality sequence stop. 391.
Location/Qualifiers
1..566
/organism="Glycine max"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-cl080-558"

```

```

FEATURES
source

```

```

TITLE
JOURNAL
COMMENT

```

```

FEATURES
source

```

```

/clone_lib="Gm-cl080"
/tissue_type="Roots of 8 day old 'Bragg' supernodulating
mutant NTS382 seedlings"
/dev_stage="8 days old"
/lab_host="DH10B"
/note="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2:
XhoI; The mRNA was isolated from roots of 8 day old
'Bragg' supernodulating mutant NTS382 seedlings that were
infected with Bradyrhizobium japonicum, strain USDA 110, 72
hours prior to harvest. Dr. Gary Stacey generously donated
the tissue. The roots were flash-frozen in liquid
nitrogen. Stratagene's cDNA Synthesis Kit (catalog number
200401) was used to synthesize the cDNA. First-strand
synthesis was performed with 5-methyl dCTP, hence the
ligated cDNA was hemimethylated. A modification of
Stratagene's first-strand synthesis primer was used. An
'anchor' nucleotide (V=A, C, or G) was added to the 3' end
of the primer [GAGAGAGAGAGAGAGAGACTGTCGAG(T18V)] to
anchor the primer at the 5' end of the poly(A) tract.
After second-strand synthesis, the cDNA ends were filled
in with cloned Pfu DNA, ligated to EcoRI adapters and
subsequently phosphorylated. The cDNA was then
precipitated and redissolved in sterile, RNase-,
DNase-free water. The XhoI site within the first- strand
synthesis primer was then restricted by digestion with
XhoI from Promega (400/ul); all XhoI sites in the cDNA
would be protected by their hemimethylated status. The
cDNA constructs were size-fractionated with a 500bp
cutoff, using Sephacryl S-500 High Resolution (Pharmacia
Biotech) in a 2-mm diameter column and a bed volume of
approximately 1ml. The column eluent was precipitated,
redissolved, and ligated into Stratagene's pBluescript II
XR Predigested vector (pBluescript II SK(+)) vector that
has been digested with EcoRI and XhoI, and phosphorylated
by Stratagene). This library was constructed in the
laboratory of Dr. Paul Keim and Dr. Virginia H. Coryell at
Northern Arizona University."
BASE COUNT      136 a      135 c      155 g      140 t
ORIGIN

```

```

Query Match      2.9%; Score 34.8; DB 13; Length 566;
Best Local Similarity 55.9%; Pred. No. 92;
Matches 66; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

QY 387 TTCTGAATGACATCAGGATCCGCCGCTCTACTGCGATACACGGCAGCGGAGACT 446
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 283 TGCACATGATCTCAGCAGGACCACTTCATGTCGTGGCGAATGTTGGCAAGCGCAAGTC 224
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 447 GAGCCTTCAGCAGTACCAACCAACGCTTCATTTCCGAAGGATGACCCACAC 504
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 223 ATGGCCTCAACCACTGTCAAAACCAAAAGCCTGTATACAGCAGGTCGCAAGAACAC 166
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

RESULT 49
BI427022/c
LOCUS
DEFINITION
sag31a10.y1 Gm-cl082 Glycine max cDNA clone SOYBEAN CLONE ID:
Gm-cl082-3884 5' similar to SW:SUSY_SOYBN P13708 SUCROSE SYNTHASE
; mRNA sequence.
BI427022
BI427022.1 GI:21678421
EST.
soybean.

```

```

ORGANISM
Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.
1 (bases 1 to 581)
Shoemaker,R., Keim,P., Vodkin,L., Erpelting,J., Coryell,V., Khanna
,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,
Wyllie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers
,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk
,R., Waterston,R. and Wilson,R.
Public Soybean EST Project
Unpublished (1999)
Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available through: Resgen, Invitrogen Corp. 2130
South Memorial Parkway Huntsville, AL 35801 For further information
call: (800)-533-4363 or contact via email: ccu@resgen.com
High quality sequence stop. 391.
Location/Qualifiers
1..581
/organism="Glycine max"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-cl080-558"

```

```

FEATURES
source

```

TITLE	JOURNAL	COMMENT
1. The Role of the Teacher in the Classroom	Journal of Educational Research	1980, Vol. 83, No. 1, pp. 1-10
2. The Impact of Technology on Education	Journal of Educational Research	1980, Vol. 83, No. 2, pp. 11-20
3. The Importance of Parental Involvement	Journal of Educational Research	1980, Vol. 83, No. 3, pp. 21-30
4. The Effect of Teacher Expectations on Student Achievement	Journal of Educational Research	1980, Vol. 83, No. 4, pp. 31-40
5. The Role of the School in the Community	Journal of Educational Research	1980, Vol. 83, No. 5, pp. 41-50
6. The Impact of Teacher Education on Classroom Practice	Journal of Educational Research	1980, Vol. 83, No. 6, pp. 51-60
7. The Importance of Teacher Collaboration	Journal of Educational Research	1980, Vol. 83, No. 7, pp. 61-70
8. The Effect of Teacher Attitudes on Student Behavior	Journal of Educational Research	1980, Vol. 83, No. 8, pp. 71-80
9. The Role of the School in the Development of the Child	Journal of Educational Research	1980, Vol. 83, No. 9, pp. 81-90
10. The Impact of Teacher Training on Student Learning	Journal of Educational Research	1980, Vol. 83, No. 10, pp. 91-100

REFERENCE
AUTHORS

```

1. .581
  /organism="Glycine max"
  /db_xref="taxon:3847"
  /clone="SOYBEAN CLONE ID: Gm-cl082-3894"
  /clone_lib="Gm-cl082"
  /tissue_types="Roots of bulked Minsoy x Noir RI progeny"
  /lab_host="DH10B"

```

TITLE	JOURNAL	COMMENT
1. The Role of the Teacher in the Classroom	Journal of Educational Research	1965, Vol. 68, No. 1, pp. 1-10
2. The Impact of Technology on Education	Journal of Educational Technology	1970, Vol. 1, No. 2, pp. 1-10
3. The Importance of Parental Involvement	Journal of Educational Research	1975, Vol. 78, No. 3, pp. 1-10
4. The Effect of Teacher Expectations on Student Achievement	Journal of Educational Research	1980, Vol. 83, No. 4, pp. 1-10
5. The Role of the School in the Community	Journal of Educational Research	1985, Vol. 88, No. 5, pp. 1-10
6. The Impact of Teacher Education on Student Achievement	Journal of Educational Research	1990, Vol. 93, No. 6, pp. 1-10
7. The Role of the Teacher in the Classroom	Journal of Educational Research	1995, Vol. 98, No. 7, pp. 1-10
8. The Impact of Technology on Education	Journal of Educational Technology	2000, Vol. 1, No. 8, pp. 1-10
9. The Importance of Parental Involvement	Journal of Educational Research	2005, Vol. 108, No. 9, pp. 1-10
10. The Effect of Teacher Expectations on Student Achievement	Journal of Educational Research	2010, Vol. 113, No. 10, pp. 1-10
11. The Role of the School in the Community	Journal of Educational Research	2015, Vol. 118, No. 11, pp. 1-10
12. The Impact of Teacher Education on Student Achievement	Journal of Educational Research	2020, Vol. 123, No. 12, pp. 1-10

BASE COUNT 128 a 138 c 154 g 161 t
ORIGIN

Query Match 2.9%; Score 34.8; DB 14; Length 581;
Best Local Similarity 55.9%; Pred. No. 92;

387	TTTCTGAATGACATCATAGGATCCCGCCGTTCTACTGCGATATACGGGCACGCCGGAGACT	446
Qy		
204	TCGCAATGATGTCACGAGGACCAATTCATCTGTGGCGAATGTTGGCAAGCGCGCAATGC	145
Db		
447	GACGCTTTCAGCCAGTACCATACCAAAAGCGCTTCATTTTCCGAAGGCATGACCAACAC	504
Qy		
144	ATGGGCTCAACCACTGTCAAAACAAAGCGCTTCGTATACAGCAGCGCTGCAGAAAGCAC	87
Db		

BI/6/12/7/C	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM
-------------	-------	------------	-----------	---------	----------	--------	----------

A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Willie, T., Underwood K., Stentoe M., Theising B., Allen M., Bowers

W. J. Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk,
P., Pitter, E., Kohn S., Shin W., Jackson V., Cardenas M., McCann

R., Waterston, R. and Wilson, R.
Public Southern ESM District

Public Soybean EST Project
Unpublished (1999)
Contact: Shoemaker R/Public Soybean EST Project

Contract: Encumbrance: 07-00102 Soybean Est Project
Public Soybean EST Project
Washington University School of Medicine

Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800

TEL: 314 286 1800
FAX: 314 286 1810
E-mail: cc@antec.com

This clone is available through: ResGen, Invitrogen Corp. 2130
Email: estewartson.wustl.edu

South Memorial Parkway Huntsville, AL 35801 For further information call: (800)-533-4363 or contact via email: ccu@resgen.com

Seq primer: -40RP from Gibco
High quality sequence stop: 413.

Location/Qualifiers
1. .584

```
/organism="Glycine max"  
/db_xref="taxon:3847"
```

```
/clone="GENOME SYSTEMS CLONE ID: Gm-cl082-2179"  
/clone_lib="Gm-cl082"
```

```
/tissue_type="Roots of bulked Minsoy x Noir RI progeny"  
/lab_host="DH10B"
```

```
/note="Vector: pBluescript II SK+; Site_1: EcoRI; Site_2:
XhoI. The mRNA was isolated from roots of bulked Minsoy x
```

Noir RI progeny that are phenotypically susceptible to flooding damage. The plants were flooded (submerged) for

24 hours prior to harvest. Dr. Tara VanToal generously provided the isolated mRNA. The roots were flash-frozen

protein and substrate mix, and 1000 µl of liquid nitrogen. Stratagene's cDNA Synthesis Kit (catalog number 200401) was used to synthesize the cDNA.

Reaction buffer 200017 was used to synthesize the cDNA. First-strand synthesis was performed with 5-methyl dCTP, hence the ligated cDNA was hemimethylated. A modification

hence the ligated cDNA was hemimethylated, a modification of Stratagene's first-strand synthesis primer was used. An /anchor/ nucleotide (5'-N-A-C or C) was added to the 3' end

of the primer [GAGAGAGAGAGAGAGAACTAGTCTCGAG(T)18V] to

anchor the primer at the 3' end of the poly(A) tract. After second-strand synthesis, the cDNA ends were filled

in with cloned pT7 DNA, ligated to EcoRI adapters and subsequently phosphorylated. The cDNA was then

precipitated and redissolved in sterile, RNase-, DNase-free water. The xhoI site within the first-strand

synthesis primer was then restricted by digestion with XhoI from Promega (40U/ul); all XhoI sites in the cDNA

would be protected by their hemimethylated status. The cDNA constructs were size-fractionated with a 500bp cutoff

, using Sephacryl S-500 High Resolution (Pharmacia Biotech) in a 2-mm diameter column and a bed volume of

approximately 1 ml. The column eluent was precipitated, redissolved and ligated into Stratagene's Bluescript II

XR Predigested vector (pBluescript II SK(+)) vector that has been digested with EcoRI and XhoI, and phosphorylated

has been targeted with books and movies, and phosphorylated

c 101 27.8 2.4 10747 5 PCT-US93-03985-1
c 102 27.8 2.4 13149 4 US-09-004-838-87
c 103 27.8 2.4 15062 3 US-09-004-838-89
c 104 27.8 2.4 15202 3 US-08-922-635-21
c 105 27.8 2.4 24979 3 US-08-147-777-3
c 106 27.8 2.4 24979 3 US-08-452-872-3
c 107 27.8 2.4 24979 5 PCT-US93-03985-3
c 108 27.8 2.4 50341 1 US-08-247-901C-1
c 109 27.8 2.4 50341 2 US-09-075-904-1
c 110 27.8 2.4 52297 4 US-09-426-436-1
c 111 27.8 2.4 52297 4 US-08-705-557-1
c 112 27.6 2.3 215 1 US-08-238-163-5
c 113 27.6 2.3 227 4 US-09-397-787-119
c 114 27.6 2.3 1001 4 US-09-641-638-553
c 115 27.6 2.3 2032 4 US-09-500-063-1
c 116 27.6 2.3 2235 4 US-09-334-601-3
c 117 27.6 2.3 2288 4 US-09-334-601-1
c 118 27.6 2.3 3243 1 US-08-611-107-32
c 119 27.6 2.3 3494 1 US-09-334-601-5
c 120 27.4 2.3 218 1 US-08-248-474-100
c 121 27.4 2.3 218 3 US-08-756-849-100
c 122 27.4 2.3 498 4 US-09-328-111-252
c 123 27.4 2.3 540 4 US-08-793-109-13
c 124 27.4 2.3 600 2 US-08-679-217-1
c 125 27.4 2.3 804 4 US-08-793-109-8
c 126 27.4 2.3 925 4 US-09-134-001C-2575
c 127 27.4 2.3 2219 3 US-09-221-017B-927
c 128 27.4 2.3 2219 3 US-08-510-646B-17
c 129 27.4 2.3 162450 4 US-09-345-882-1
c 130 27.2 2.3 1374 1 US-08-123-761A-2
c 131 27.2 2.3 1455 4 US-09-615-192A-245
c 132 27.2 2.3 1956 3 US-08-693-940-2
c 133 27.2 2.3 1956 4 US-09-566-660-2
c 134 27.2 2.3 2266 1 US-08-453-472-1
c 135 27.2 2.3 2266 1 US-08-453-952-1
c 136 27.2 2.3 2266 2 US-08-484-993B-42
c 137 27.2 2.3 2266 2 US-08-862-903-1
c 138 27.2 2.3 2266 2 US-08-484-159B-42
c 139 27.2 2.3 2266 2 US-08-484-596A-42
c 140 27.2 2.3 2266 2 US-08-480-150A-42
c 141 27.2 2.3 2266 3 US-08-458-731-42
c 142 27.2 2.3 2266 3 US-08-149-223A-42
c 143 27.2 2.3 2291 1 US-07-872-644-5
c 144 27.2 2.3 2291 1 US-08-297-494-5
c 145 27.2 2.3 2291 1 US-08-297-510-5
c 146 27.2 2.3 2291 1 US-08-479-532-5
c 147 27.2 2.3 2291 1 US-08-455-526-5
c 148 27.2 2.3 2291 1 US-08-455-526-5
c 149 27.2 2.3 2291 3 US-09-139-491-5
c 150 27.2 2.3 2291 5 PCT-US92-03222-5

ALIGNMENTS

RESULT 1
US-09-103-840A-2
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103, 840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765

; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2
Query Match 4.8%; Score 57.2; DB 4; Length 4403765;
Best Local Similarity 54.2%; Pred. No. 1.le-06;
Matches 116; Conservative 0; Mismatches 98; Indels 0; Gaps 0;
QY 272 TTCTGTGGCGTGGCTGCATCATCCGGAAGCTTCCGTCGGGTAAATAAATCGCGCAG 331
Db 2447373 TTCTGTGGCGGCGGCCCATCGGACGCGCGGATCCGATGATCAGCACTCGGCCAC 2447432
QY 332 TGGCGCGGTCCATGACAGACATCCCGGACGCGGTACACGCTCCCTGTACATTTCTGTG 391
Db 2447433 GCGTGTGGCGCACCGGTCACCGACCTACCGTGGACCACTAGCCAGTCTTGTGTGCTG 2447492
QY 392 AATGACATCAGGATCCCGCGCTCTCACTGGCGATACGGGCACGCCGAGACTGACGC 451
Db 2447493 CACCGTTTCGGCGCTCCGCGCAATTCGCGCGGATTACCGCACGCCGCGGAGGC 2447552
QY 452 TTCAGCAGTACCATACCAACCGCTTCATTTCC 485
Db 2447553 TTCAGGACACGATGCCCAAGCCCTCGACGTCC 2447586
RESULT 2
US-09-103-840A-1
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103, 840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37RV
US-09-103-840A-1
Query Match 4.8%; Score 57.2; DB 4; Length 4411529;
Best Local Similarity 54.2%; Pred. No. 1.le-06;
Matches 116; Conservative 0; Mismatches 98; Indels 0; Gaps 0;
QY 272 TTCTGTGGCGTGGCTGCATCATCCGGAAGCTTCCGTCGGGTAAATAAATCGCGCAG 331
Db 2450072 TTCTGTGGCGGCGGCCCATCGGACGCGGCCGATCCGATGATCAGCACTCGGCCAC 2450131
QY 332 TGGCGCGGTCCATGACAGACATCCCGGACGCGGTAAACAGGCTCCCTGTACATTTCTGTG 391
Db 2450132 GCGTGTGGCGCACCGGTCACCGACCTACCGTGGACCACTAGCCAGTCTTGTGTGCTG 2450191
QY 392 AATGACATCAGGATCCCGCGCTCTCACTGGCGATACGGGCACGCCGAGACTGACGC 451
Db 2450192 CACCGTTTCGGCGCTCCGCGCAATTCGCGCGGATTACCGCACGCCGCGGAGGC 2450251
QY 452 TTCAGCAGTACCATACCAACCGCTTCATTTCC 485
Db 2450252 TTCAGGACACGATGCCCAAGCCCTCGACGTCC 2450285

RESULT 3

US-09-103-840A-2/c
 ; Sequence 2, Application US/09103840A
 ; Patent No. 6294328
 ; GENERAL INFORMATION:
 ; APPLICANT: FLEISCHMAN, Robert D.
 ; APPLICANT: WHITE, Owen R.
 ; APPLICANT: FRASER, Claire M.
 ; APPLICANT: VENTER, John C.
 ; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
 ; FILE OF INVENTION: TUBERCULOSIS
 ; FILE REFERENCE: 24366-20007.00
 ; CURRENT APPLICATION NUMBER: US/09/103,840A
 ; CURRENT FILING DATE: 1998-06-24
 ; NUMBER OF SEQ ID NOS: 2
 ; SOFTWARE: Patentin ver. 2.1
 ; SEQ ID NO 2
 ; LENGTH: 4403765
 ; TYPE: DNA
 ; ORGANISM: Mycobacterium tuberculosis
 ; FEATURE:
 ; OTHER INFORMATION: CDC 1551
 ; OTHER INFORMATION: "n" bases at various positions throughout the sequence
 ; OTHER INFORMATION: represent a, t, c or g
 US-09-103-840A-2

Query Match 3.4%; Score 39.8; DB 4; Length 4403765;
 Best Local Similarity 50.8%; Pred. No. 0.65;
 Matches 95; Conservative 0; Mismatches 92; Indels 0; Gaps 0;
 QY 333 GCGCGGTCTCATGACACATCCCCACGGGTAACAGCGTCCCTGTTCATCTTCTGA 392
 DB 577989 GCGTCGGCCCACTGACCGACTCGTCCCGGACACCGAGGTGCGGTGATCCCGTCCGC 577930
 QY 393 ATGACATCAGGGATCCCGCCGCTCTCACTGGCGATACGGGCGAGACTGACGCT 452
 DB 577929 ACCGCGAGGGCGCGCCCGCCCGCGCGCCACACCGGTGTGCGCGAGCTTGGGCC 577870
 QY 453 TCAGCCAGTACCATACCAACAGCTTCAATTTCCGAAGGCGATGACCAACACACTGGCAATC 512
 DB 577869 TCCACAGCAACAGCGCGGACGACTCGGAGTAGTCTCGGACCGACCGACAGGTCCGCGGCC 577810
 QY 513 CGGTAGA 519
 DB 577809 CGAAACA 577803

RESULT 4

US-09-103-840A-1/c
 ; Sequence 1, Application US/09103840A
 ; Patent No. 6294328
 ; GENERAL INFORMATION:
 ; APPLICANT: FLEISCHMAN, Robert D.
 ; APPLICANT: WHITE, Owen R.
 ; APPLICANT: FRASER, Claire M.
 ; APPLICANT: VENTER, John C.
 ; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
 ; FILE OF INVENTION: TUBERCULOSIS
 ; FILE REFERENCE: 24366-20007.00
 ; CURRENT APPLICATION NUMBER: US/09/103,840A
 ; CURRENT FILING DATE: 1998-06-24
 ; NUMBER OF SEQ ID NOS: 2
 ; SOFTWARE: Patentin ver. 2.1
 ; SEQ ID NO 1
 ; LENGTH: 4411529
 ; TYPE: DNA
 ; ORGANISM: Mycobacterium tuberculosis
 ; OTHER INFORMATION: H37Rv
 US-09-103-840A-1
 Query Match 3.4%; Score 39.8; DB 4; Length 4411529;
 Best Local Similarity 50.8%; Pred. No. 0.65;
 Matches 95; Conservative 0; Mismatches 92; Indels 0; Gaps 0;

QY 333 GCGCGGTCTCATGACACATCCCCACGGGTAACAGCGTCCCTGTTCATCTTCTGA 392
 DB 576547 GCGTCGGCCCACTGACCGACTCGTCCCGGACACCGAGGTGCGGTGATCCCGTCCGC 576488
 QY 393 ATGACATCAGGGATCCCGCCGCTCTCACTGGCGATACGGGCGAGACTGACGCT 452
 DB 576487 ACCGCGAGGGCGAGCCCGCCCGCCCGCGCGCCACCGCGGTGTGCGCGAGCTTGGGCC 576428
 QY 453 TCAGCCAGTACCATACCAACAGCTTCAATTTCCGAAGGCGATGACCAACACACTGGCAATC 512
 DB 576427 TCCACAGCAACAGCGCGGACGACTCGGAGTAGTCTCGGACCGACCGACAGGTCCGCGGCC 576368
 QY 513 CGGTAGA 519
 DB 576367 CGAAACA 576361

RESULT 5

5405943-3/c
 ; Patent No. 5405943
 ; APPLICANT: COMINGS, DAVID E.
 ; TITLE OF INVENTION: TOURETTE SYNDROM, AUTISM AND ASSOCIATED
 ; BEHAVIORS
 ; NUMBER OF SEQUENCES: 5
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/07/562,596
 ; FILING DATE: 03-AUG-1990
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 125,577
 ; FILING DATE: 25-NOV-1987
 ; APPLICATION NUMBER: 271,653
 ; FILING DATE: 16-NOV-1988
 ; APPLICATION NUMBER: 410,831
 ; FILING DATE: 22-SEP-1989
 ; SEQ ID NO:3
 ; LENGTH: 1647
 5405943-3

Query Match 2.9%; Score 33.8; DB 6; Length 1647;
 Best Local Similarity 53.4%; Pred. No. 0.63;
 Matches 71; Conservative 0; Mismatches 62; Indels 0; Gaps 0;
 QY 1013 ATTACTGTATGCGCACTATGGCAGATGACGAGTTAAATTCCTGATTAATTCG 1072
 DB 345 ATCTCTGACATGGCCATCTCTGAAAGATCTCTGAAACAGAAATCCAACTCCAGAGATTG 286
 QY 1073 TCGAAGTCTGAGGATGGAAGGAGTGAAGGCTCTTCTGAAAGGATAAAAGTGACATC 1132
 DB 285 CTTAACCAGAGTTCAATGAGTTATGATAAAAGATGTTTCATCATGATGATTTT 226
 QY 1133 ATGCCCTCTTTT 1145
 DB 225 ATTTCCTTTTGT 213

RESULT 6

US-09-066-046-1/c
 ; Sequence 1, Application US/09066046A
 ; Patent No. 6204252
 ; GENERAL INFORMATION:
 ; APPLICANT: MURPHY, Cheryl
 ; STOREY, James
 ; BELTZ, Gerald A.
 ; COUGHLIN, Richard T.
 ; TITLE OF INVENTION: CHARACTERIZATION OF GRANULOCYTIC
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESS: HALE AND DORR LLP
 ; STREET: 60 State Street
 ; CITY: Boston
 ; STATE: MA
 ; NUMBER OF SEQUENCES: 65
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESS: HALE AND DORR LLP
 ; STREET: 60 State Street
 ; CITY: Boston
 ; STATE: MA

COUNTRY: United States
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/066,046A
FILING DATE: 24-Apr-1998
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Superko, Colleen
REGISTRATION NUMBER: 39,850
REFERENCE/DOCKET NUMBER: 106,941,155
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 526-6000
TELEFAX: (617) 526-6000
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3147 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-066-046-1

Query Match 2.8%; Score 33; DB 4; Length 3147;
Best Local Similarity 48.2%; Pred. No. 1.7;
Matches 93; Conservative 0; Mismatches 100; Indels 0; Gaps 0;

QY 397 CATCAGGATCCGCCGCTCTCACTGGCGATACGGGCACCGGAGACTGACGCTTCAG 456
DB 1253 CAACAGTATATCTCTCAGTAGTATCTTCGAGCAGGTAACGTCAGTGTGGCGATTCTG 1194

QY 457 CCAGTACCATCAACCAACGCTTCATTTCCGAAGGCATGACACACATCGGCAATCCGGT 516
DB 1193 GCACATAATGATGATCTACTTCTAGTCAGTAGTGTGGTTAACAACCGCAT 1134

QY 517 AGACCGGTAACTGGGAAAGGCGACCTGCCATTACACATCTCCGCTCATTCGCCAGGT 576
DB 1133 TTGAGGCTCTGTTGCTATAGGGCAGTAGACATAAGTACAGATGGTTGCACCTCCACTG 1074

QY 577 GTTCTGCTGCTG 589
DB 1073 ATTCTTGTGCTG 1061

RESULT 7
US-09-066-046-7/c
Sequence 7, Application US/09066046A
Patent No. 6204252
GENERAL INFORMATION:
APPLICANT: MURPHY, Cheryl
STOREY, James
BELTZ, Gerald A.
COUGHLIN, Richard T.
TITLE OF INVENTION: CHARACTERIZATION OF GRANULOCYTIC
NUMBER OF SEQUENCES: 65
CORRESPONDENCE ADDRESS:
ADDRESSEE: HALE AND DORR LLP
STREET: 60 State Street
CITY: Boston
STATE: MA
COUNTRY: United States
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/066,046A
FILING DATE: 24-Apr-1998
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Superko, Colleen
REGISTRATION NUMBER: 39,850
REFERENCE/DOCKET NUMBER: 106,941,155
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 526-6000
TELEFAX: (617) 526-6000
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 5570 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-09-066-046-7

Query Match 2.8%; Score 33; DB 4; Length 5570;
Best Local Similarity 48.2%; Pred. No. 2.3;
Matches 93; Conservative 0; Mismatches 100; Indels 0; Gaps 0;

QY 397 CATCAGGATCCGCCGCTCTCACTGGCGATACGGGCACCGGAGACTGACGCTTCAG 456
DB 1007 CAACAGTATATCTCTCAGTAGTATCTTCGAGCAGGTAACGTCAGTGTGGCGATTCTG 948

QY 457 CCAGTACCATCAACCAACGCTTCATTTCCGAAGGCATGACACACATCGGCAATCCGGT 516
DB 947 GCACATAATGATGATCTACTTCTAGTCAGTAGTGTGGTTAACAACCGCAT 888

QY 517 AGACCGGTAACTGGGAAAGGCGACCTGCCATTACACATCTCCGCTCATTCGCCAGGT 576
DB 887 TTGAGGCTCTGTTGCTATAGGGCAGTAGACATAAGTACAGATGGTTGCACCTCCACTG 828

QY 577 GTTCTGCTGCTG 589
DB 827 ATTCTTGTGCTG 815

RESULT 8
US-09-141-000-2
Sequence 2, Application US/09141000
Patent No. 6054295
GENERAL INFORMATION:
APPLICANT: Chen, Fang
TITLE OF INVENTION: DNA MOLECULES ENCODING HUMAN NUCLEAR
FILE REFERENCE: 19999Y
CURRENT APPLICATION NUMBER: US/09/141,000
CURRENT FILING DATE: 1998-08-26
NUMBER OF SEQ ID NOS: 30
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 2
LENGTH: 500
TYPE: PRT
ORGANISM: Human
US-09-141-000-2

Query Match 2.8%; Score 32.6; DB 3; Length 500;
Best Local Similarity 10.8%; Pred. No. 0.78; Indels 0; Gaps 0;
Matches 43; Conservative 110; Mismatches 246;

QY 733 GCGGAACATTGCTGCTGATGTCAGCCAGGCGGTGATGAGGGGTAAAGCGTCGATGG 792
DB 24 SSGIDALSHHSPSGSSDASGGFGIALGTHANGLDSPPMFAGAGLGGTPCRKSYEDCAGI 83

```

QY 793 CTTCAATATCAATCGGGCCAGTCGAAACCCGGTGAATAACGGTTACCGGTCTCCTGA 852
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
84 MEDSAIKCEYMLNAIPRLCLVCGDIASGHYGVASCEACKAFKFTIOGNIYSPATN 143
QY 853 CACCTTCGCCCATCAGATGCGCCATCATCGGTGAGATAGACACAAATGAATCACA 912
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
144 ECEITRRRRKSCQACRFMKCLKVGMLKEGVRDLRVRGRGKRYKRLDSESSPYLSQISP 203
QY 913 GATAATCAGGGAAGCTTCTCGTCTTACGGGTGATGATAGTTTCTGCTGACAAATG 972
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
204 PAKPLTKIVSYLLVAEPDKLYAMPMPPEGDIKALTTLCDLADRELVIIGWAKHIPG 263
QY 973 TGAAGCGGTGACAGCATATCAGACGGCTGCTGCTATATCTGCTATCTGCTGCGCCAT 1032
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
264 FSSLSLGDQMSLQSAWELIIGIVYRSLPYDDKLVIADYIMDEHSHSLAGLLELYRA 323
QY 1033 GCGAGATGACGATCAGGTGTTAAATCCCGGATATCCGCGAAGTCTGAGGATGAAG 1092
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
324 ILQLVRRYKRLKVEKEEFVTLKALALANSDSMYIEDLEAVQKLQDLLHEALQDYELSQRH 383
QY 1093 GAAGTGAAGCGTGTCTCGTGAAGGAATAAAGTGACAT 1131
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
384 EEPWRTGKLLTLTLLRLQTAAKAVOHFYSVKLQGVPMH 422

```

RESULT 9

```

US-09-028-327-2/c
; Sequence 2, Application US/09028327
; Patent No. 6130064
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: HUMAN SMN-LIKE PROTEIN
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Dr.
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/028,327
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0482 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2426 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: BRSTNOT24
; CLONE: 3769729
; US-09-028-327-2

```

```

Query Match 2.88; Score 32.6; DB 3; Length 2426;
Best Local Similarity 54.68; Pred. No. 2;
Matches 65; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

QY 800 ATGCAATCGGGCCAGTCGAAACCCGGTGAATAACGGTTACCGGTCTCCTGACACCTTC 859
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
254 AGCGGTTTCGGGATCGAGAGACCCACTGTAAATATATGACACAAGTGACTTGCACCCCTGC 195
QY 860 CGCATCATGATGCGCCATCATCGGTGAGATAGGACACAAATGAATCACAACAGATAAT 918
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
194 CTCCTTTAGAACAGCTTATTCAGCAAAACAGCCTAATAATTGACTATCTCTGATTAT 136

RESULT 10
US-08-035-558-1/c
; Sequence 1, Application US/08035558
; Patent No. 5462734
; GENERAL INFORMATION:
; APPLICANT: Letchworth, Geoffrey J.
; APPLICANT: Israel, Barbara A.
; TITLE OF INVENTION: HERPESVIRUS VACCINE AND METHOD
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Quarles and Brady
; STREET: P.O. BOX 2113
; CITY: MADISON
; STATE: WISCONSIN
; COUNTRY: U.S.A.
; ZIP: 53701
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette - 3.50 inch, 800Kb storage
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Macintosh
; SOFTWARE: Microsoft Word 4.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/035,558
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/607,794
; FILING DATE: 02-NOV-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Nicholas J. Seay
; REGISTRATION NUMBER: 27,386
; REFERENCE/DOCKET NUMBER: 9629691682
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (608)251-5000
; TELEFAX: (608)251-9166
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3519 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: genomic DNA
; HYPOTHETICAL: no
; ANTI-SENSE: no
; ORIGINAL SOURCE:
; ORGANISM: bovine herpesvirus-1
; INDIVIDUAL ISOLATE: P8-2 strain
; FEATURE:
; NAME/KEY: coding sequence for g1 glycoprotein
; LOCATION: 379 to 3165
; IDENTIFICATION METHOD: sequence analysis
; PUBLICATION INFORMATION:
; AUTHORS: Misra, Vikram
; AUTHORS: Nelson, Randy
; AUTHORS: Smith, Michael
; TITLE: Sequence of a Bovine Herpesvirus Type-1 etc.
; JOURNAL: Virology
; VOLUME: 166

```

PAGES: 542-549
DATE: 1988
RELEVANT RESIDUES IN SEQ ID NO: 1: FROM 379 TO 3165
US-08-035-558-1

Query Match 2.7%; Score 32.2; DB 1; Length 3519;
Best Local Similarity 49.48; Pred. No. 3.3; Indels 0; Gaps 0;
Matches 82; Conservative 0; Mismatches 84; Indels 0; Gaps 0;
QY 296 CCGAAGCGTTCCGGTCGGGATAAAATCCGCGAGTCGCCCGGTCATCGACACATC 355
DB 196 CCGNGCCCGCGGATCGGAAGCAAAAAAATGCTGAAGTCGTTACCGCCCGCCGCC 137
QY 356 CCCACGGGTAAACAGCGTCCCTGTGCATCTTCTGAATGACATCAGGATCCGCCCGT 415
DB 136 CCGCGCGCCCGCAGCAATAACGCGCCAGTCCTCTTGAGGTCCGGGAGCGAGCCAC 77
QY 416 CTCAGTCGCGATAACGGCGACCGCGGAGACTGACGCTTCAGCCAGT 461
DB 76 GTTTCACGCTAAAGTACAGCGCGGTGTTTGGCGCTGCGGAGT 31

RESULT 11
US-08-232-463-14/c
Sequence 14, Application US/08232463
Patent No. 5670367
GENERAL INFORMATION:
APPLICANT: DORNER, F.
APPLICANT: SCHEIFLINGER, F.
APPLICANT: FALKNER, F. G.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: ptz9pt-Fls
US-08-232-463-14

Query Match 2.7%; Score 32.2; DB 1; Length 7218;
Best Local Similarity 5.38; Pred. No. 5;
US-08-232-463-14

Matches 22; Conservative 205; Mismatches 188; Indels 0; Gaps 0;
QY 759 CAGGCGTCATATGACGGTAAACGTCGCGATGGCTTCATTATGCAATGCGGCCAGTCG 818
DB 1435 CRR 1376
QY 819 AAACCGGTGAATAACCGTTACCGGTGTCTTGACACCTTCGCGCATCAGATCGCCCATC 878
DB 1375 RRR 1316
QY 879 ATGGGTGAGATAGGACACAATAATGAAATCAGACACATAATTCAGGAAACGTTCTGTC 938
DB 1315 RRR 1256
QY 939 TTACGGGTGATGTAGTGTCTGACAAATGTAAGCGGTGACAGCATATCAGACGG 998
DB 1255 RRR 1196
QY 999 CTCAGTCCTGCTATATGTCATGCGCCACTATGCGCAGATGACGATCAGTTAAAT 1058
DB 1195 RRR 1136
QY 1059 TCCCGCATATCGTCTGAAGTCTGAGGATGGAAGGAGTGAAGGCTGTCTCTGAAAGA 1118
DB 1135 RRR 1076
QY 1119 ATAAAGTGACATCATCCCTCTTTTCTGGCTTCGCGAGCAATTTACTTTTTT 1173
DB 1075 RRR 1021

RESULT 12
US-08-232-463-14
Sequence 14, Application US/08232463
Patent No. 5670367
GENERAL INFORMATION:
APPLICANT: DORNER, F.
APPLICANT: SCHEIFLINGER, F.
APPLICANT: FALKNER, F. G.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs


```

; LENGTH: 3031 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; STRAIN: Shc Proteins
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 82..1503
; US-08-664-962B-3

```

```

Query Match 2.6%; Score 31.2; DB 4; Length 3031;
Best Local Similarity 54.3%; Pred. No. 6.5;
Matches 63; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

QY 358 CCACGGGTAAACAGCGTCCCTGTACACATTCTTGAATGACATCAGGATCCCGCGTCT 417
Db 591 CCTGTGAATCAGAGAGCTGACACATTCTGGAGTCTCCGAGGCGTTCGCCAGGATGT 650

QY 418 CACTGGCGATACAGCGGACGCGGAGACTGACGCTTCAGCCAGTACCAACCAAC 473
Db 651 CATCAGCACCATTGGCCAGGCTTCGAGTTGCGCTTCAAAACAATACCTCAGGAACC 706

```

```

RESULT 18
US-09-311-743-3
; Sequence 3, Application US/09311743
; Patent No. 6238903
; GENERAL INFORMATION:
; APPLICANT: Krystal, Gerald
; TITLE OF INVENTION: SH2-CONTAINING INOSITOL-PHOSPHATASE
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BERESKIN & PARR
; STREET: 40 King Street West
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5H 3Y2
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/311,743
; FILING DATE: 14-May-1999
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Gravelle, Micheline
; REGISTRATION NUMBER: 40,261
; REFERENCE/DOCKET NUMBER: 7771-32
; TELEPHONE: 416-364-7311
; TELEFAX: 416-361-1398
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3031 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; STRAIN: Shc Proteins
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 82..1503
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-311-743-3

```

```

Query Match 2.6%; Score 31.2; DB 4; Length 3031;
Best Local Similarity 54.3%; Pred. No. 6.5;
Matches 63; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

QY 358 CCACGGGTAAACAGCGTCCCTGTACACATTCTTGAATGACATCAGGATCCCGCGTCT 417
Db 591 CCTGTGAATCAGAGAGCTGACACATTCTGGAGTCTCCGAGGCGTTCGCCAGGATGT 650

QY 418 CACTGGCGATACAGCGGACGCGGAGACTGACGCTTCAGCCAGTACCAACCAAC 473
Db 651 CATCAGCACCATTGGCCAGGCTTCGAGTTGCGCTTCAAAACAATACCTCAGGAACC 706

```

```

RESULT 19
PCT-US95-03747-1/c
; Sequence 1, Application PC/TUS9503747
; GENERAL INFORMATION:
; APPLICANT: LA JOLLA CANCER RESEARCH FOUNDATION
; TITLE OF INVENTION: Brevican, A Glial Cell Proteoglycan
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell and Flores
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/03747
; FILING DATE: 27-MAR-1995
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Imbra, Richard J.
; REGISTRATION NUMBER: 37,643
; REFERENCE/DOCKET NUMBER: FP-LJ 1453
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3259 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 112..2848
; PCT-US95-03747-1

```

```

Query Match 2.6%; Score 31.2; DB 5; Length 3259;
Best Local Similarity 51.4%; Pred. No. 6.8;
Matches 72; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

QY 408 CGCCCGCTCTCACTGGCGATAACGGGCGACGCGGAGACTGACGCTTCAGCCAGTACCAT 467
Db 2024 CTGGCACTGTCACTGGGCGAGCACTGGTGGGCGACGCTGCTGGGAGAGTTCCTGGGAG 1965

QY 468 CCAACGCTTCATTTTCCGAAGGATGACACACACACTGCGCAATCCGGTAGACCGGTAA 527
Db 1964 CTTCTGGAATCTCTTCAGAGGGGGTCTCCAGATCCCTGGTATCCCGGCGCGGTG 1905

QY 528 GCTGGGAAAGGGCACCTGC 547
Db 1904 GCTGGAAGCAGGGAAGGGC 1885

```

RESULT 20
US-08-896-449A-1/c

Sequence 1, Application US/08896449A
Patent No. 6040143
GENERAL INFORMATION:
APPLICANT: Venta, Patrick J
APPLICANT: Yuzbasiyan-Gurkan, Vilma
APPLICANT: Schall, William D
APPLICANT: Brewer, George J
TITLE OF INVENTION: DNA ENCODING CANINE VON WILLEBRAND
TITLE OF INVENTION: FACTOR AND METHODS OF USE
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Harness, Dickey & Pierce, P.L.C.
CITY: Troy
STATE: Michigan
COUNTRY: USA
ZIP: 48098
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/896,449A
FILING DATE: 18-JUL-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Smith, Deann F.
REFERENCE/DOCKET NUMBER: 2115-001226
TELECOMMUNICATION INFORMATION:
TELEPHONE: 248-641-1600
TELEFAX: 248-641-0270
TELEX: 287637
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 8802 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 203..8641
OTHER INFORMATION: /function= "Blood Clotting Protein"
OTHER INFORMATION: /product= "Canine von Willebrand Factor"
OTHER INFORMATION: /standard_name= "vWF"
PUBLICATION INFORMATION:
AUTHORS: Venta, Patrick J.
AUTHORS: Li, Jianping
AUTHORS: Yuzbasiyan-Gurkan, Vilma
AUTHORS: Schall, William D.
AUTHORS: Brewer, George J.
TITLE: Von Willebrand's Disease in the Scottish
TITLE: Terrier is Caused by a Single Base Deletion in
TITLE: Exon Four of the von Willebrand Factor Gene
JOURNAL: Journal of the American Veterinary Medicine Association
DATE: 1996
RELEVANT RESIDUES IN SEQ ID NO: 1: FROM 1 TO 8802
US-08-896-449A-1
Query Match 2.6%; Score 31.2; DB 3; Length 8802;
Best Local Similarity 50.0%; Pred. No. 12;
Matches 78; Conservative 0; Mismatches 78; Indels 0; Gaps 0;
QY 390 TGAATGACATCAGGATCCCGCGTCTCACTGCGATACGGCGCAGCGGAGACTGAC 449
DB 7659 TGAAGGACATAAGTGAAGCGCTGACAGGAGTTGTCTCAGAGGGCTTCTGGGAGCACTGG 7600
QY 450 GCTTCAGCGAGTACCATACCAACGCTTCATTTCCGAAGGCGATGACACACCACTGGCA 509
DB 7599 GCCACACGAGCCCATCACAGAGTCTCCAAAGTCCGTGCGAGGTGCACACGTACAGGCC 7540
US-08-896-449A-1
Query Match 2.6%; Score 31.2; DB 3; Length 8802;
Best Local Similarity 50.0%; Pred. No. 12;
Matches 78; Conservative 0; Mismatches 78; Indels 0; Gaps 0;
QY 390 TGAATGACATCAGGATCCCGCGTCTCACTGCGATACGGCGCAGCGGAGACTGAC 449
DB 7659 TGAAGGACATAAGTGAAGCGCTGACAGGAGTTGTCTCAGAGGGCTTCTGGGAGCACTGG 7600
QY 450 GCTTCAGCGAGTACCATACCAACGCTTCATTTCCGAAGGCGATGACACACCACTGGCA 509
DB 7599 GCCACACGAGCCCATCACAGAGTCTCCAAAGTCCGTGCGAGGTGCACACGTACAGGCC 7540

QY 510 ATCCGGTAGACCGGTAAACGCTGGGAAAGGCGACCT 545
DB 7539 TCCTCCCAAGAACTGGCCACAGAGGTAGATGGTGCCT 7504
RESULT 21
US-09-132-652-1/c
Sequence 1, Application US/09132652
Patent No. 6074832
GENERAL INFORMATION:
APPLICANT: Venta, Patrick J
APPLICANT: Yuzbasiyan-Gurkan, Vilma
APPLICANT: Schall, William D
APPLICANT: Brewer, George J
APPLICANT: Duffendeck, John
TITLE OF INVENTION: DNA ENCODING CANINE VON WILLEBRAND FACTOR AND METHODS
TITLE OF INVENTION: OF USE
FILE REFERENCE: 2115S-001226CPB
CURRENT APPLICATION NUMBER: US/09/132,652
CURRENT FILING DATE: 1998-08-11
EARLIER APPLICATION NUMBER: 08/896,449
EARLIER FILING DATE: 1997-07-18
NUMBER OF SEQ ID NOS: 29
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 1
LENGTH: 8802
TYPE: DNA
ORGANISM: Canis familiaris
US-09-132-652-1
Query Match 2.6%; Score 31.2; DB 3; Length 8802;
Best Local Similarity 50.0%; Pred. No. 12;
Matches 78; Conservative 0; Mismatches 78; Indels 0; Gaps 0;
QY 390 TGAATGACATCAGGATCCCGCGTCTCACTGCGATACGGCGCAGCGGAGACTGAC 449
DB 7659 TGAAGGACATAAGTGAAGCGCTGACAGGAGTTGTCTCAGAGGGCTTCTGGGAGCACTGG 7600
QY 450 GCTTCAGCGAGTACCATACCAACGCTTCATTTCCGAAGGCGATGACACACCACTGGCA 509
DB 7599 GCCACACGAGCCCATCACAGAGTCTCCAAAGTCCGTGCGAGGTGCACACGTACAGGCC 7540
QY 510 ATCCGGTAGACCGGTAAACGCTGGGAAAGGCGACCT 545
DB 7539 TCCTCCCAAGAACTGGCCACAGAGGTAGATGGTGCCT 7504
RESULT 22
US-09-403-768-7/c
Sequence 7, Application US/09403768
Patent No. 6444804
GENERAL INFORMATION:
APPLICANT: Lam, Joseph S.
APPLICANT: De Kievit, Teresa R.
APPLICANT: Burrows, Lori L.
APPLICANT: Mathew, Andrew
APPLICANT: Mathew, Mauricia
TITLE OF INVENTION: No. 6444804el Proteins Involved in the Synthesis and Assembly
TITLE OF INVENTION: of Core Lipopolysaccharide of Pseudomonas Aeruginosa
FILE REFERENCE: 6580-177
CURRENT APPLICATION NUMBER: US/09/403,768
CURRENT FILING DATE: 1999-11-02
PRIOR APPLICATION NUMBER: U.S. 60/045,418
PRIOR FILING DATE: 1997-05-02
PRIOR APPLICATION NUMBER: U.S. 60/046,149
PRIOR FILING DATE: 1997-05-09
NUMBER OF SEQ ID NOS: 8
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 7
LENGTH: 1122
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa rfag (waaG)

```

: FEATURE:
: OTHER INFORMATION: Description of Organism: Pseudomonas aeruginosa
: OTHER INFORMATION: serotype O5 strain PAO1
US-09-403-768-7

Query Match      2.6%; Score 30.8; DB 4; Length 1122;
Best Local Similarity 50.7%; Pred. No. 4.9;
Matches 74; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

QY 363 GGTACACGCGTCCCTGTACATTTCTTGAATGACATCAGGATCCGCCGCTCACTG 422
Db 956 GGCAGCACCGCGCGCGCTCGGCGATAGTGGCGATAGCGGACATCGGTC 897

QY 423 GCATACGCGCGCGCGGAGACTGAGCTTACGCTTACGCTACCATACCAAGCTTATTT 482
Db 896 ACCAACACCGCGCGCGGAGAGACGCGCTCCAGCAGCAGCAGCTACCGGTCTCTCGTTG 837

QY 483 TCCGAAGGCATGACCCACACACTGGC 508
Db 836 TAGCGCGGTGGATCAGCAGGTGGC 811

RESULT 23
US-08-961-527-155
: Sequence 155, Application US/08961527
: Patent No. 6420135
: GENERAL INFORMATION:
: APPLICANT: Charles Kunsch
: TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
: NUMBER OF SEQUENCES: 391
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Human Genome Sciences, Inc.
: STREET: 9410 Key West Avenue
: CITY: Rockville
: STATE: Maryland
: COUNTRY: USA
: ZIP: 20850
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
: OPERATING SYSTEM: HP Vectra 486/33
: SOFTWARE: MS-DOS version 6.2
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/961,527
: FILING DATE:
: CLASSIFICATION: 424
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER:
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Brookes, A. Anders
: REGISTRATION NUMBER: 36,373
: REFERENCE/DOCKET NUMBER: PB340P1
: TELEPHONE: (301) 309-8504
: TELEFAX: (301) 309-8512
: INFORMATION FOR SEQ ID NO: 155:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 6474 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
US-08-961-527-155

Query Match      2.6%; Score 30.8; DB 4; Length 6474;
Best Local Similarity 47.4%; Pred. No. 14;
Matches 92; Conservative 0; Mismatches 102; Indels 0; Gaps 0;

QY 351 ACATCCCCCAGGGTAACACGCGTCCCTGTGCATCTTCTTGAATGACATCAGGATCCCG 410
Db 5601 ACATCTCCCTCTCAATAAAATCCATCAACCCATGTTCAATAATTTCACTTAACCCG 5660

QY 411 CCGCTCTACTGGGATACGGGCGACCGGAGACTGACCGCTTCAGCCAGTACCATACCA 470

```

```

Db 5661 CCAACATTACTAGCTAAACCGGAGTTCTTGTGACATTGACTCTAAACACACATAGGT 5720
QY 471 AAGCGTTCAATTTCCGAAGGCATGACCCACACACTGGCAATCCGTTAGACCGGTAACGCT 530
Db 5721 ATTCTTCTCTATCAGAGGAATATACATAAATCCGATATTTGGTAAACTATAGTAGCT 5780
QY 531 GGGAAAGGGCAC 544
Db 5781 GGATAGATTTCACC 5794

RESULT 24
US-07-737-736B-5/c
: Sequence 5, Application US/07737736B
: Patent No. 5260199
: GENERAL INFORMATION:
: APPLICANT: DeLuca, Hector F.
: APPLICANT: Ross, Troy K.
: APPLICANT: Frahl, Jean M.
: TITLE OF INVENTION: Method Of Producing
: TITLE OF INVENTION: 1,25-Dihydroxyvitamin D3 Receptor Protein
: NUMBER OF SEQUENCES: 6
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Carl R. Schwartz, Esq., c/o Quarles & Brady
: STREET: 411 East Wisconsin Avenue
: CITY: Milwaukee
: STATE: Wisconsin
: COUNTRY: U.S.A.
: ZIP: 53202
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA: US/07/737,736B
: APPLICATION NUMBER: 19910730
: FILING DATE: 19910730
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Schwartz, Carl R.
: REGISTRATION NUMBER: 29,437
: REFERENCE/DOCKET NUMBER: 96-296-2185-2
: TELEPHONE: 414-277-5715
: TELEFAX: 414-277-5774
: INFORMATION FOR SEQ ID NO: 5:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1399 base pairs
: TYPE: NUCLEIC ACID
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: HYPOTHETICAL: NO
: ANTI-SENSE: NO
: ORIGINAL SOURCE:
: ORGANISM: Homo sapiens
: PUBLICATION INFORMATION:
: AUTHORS: Baker, Andrew R.
: AUTHORS: McDonnell, Donald P.
: AUTHORS: Hughes, Mark
: AUTHORS: Crisp, Tracey M.
: AUTHORS: Mangelsdorf, David J.
: AUTHORS: Haussler, Mark R.
: AUTHORS: Pike, J. W.
: AUTHORS: Shine, John
: AUTHORS: O'Malley, Bert W.
: TITLE: Cloning and expression of full-length cDNA
: TITLE: encoding human vitamin D receptor
: JOURNAL: Proc. Natl. Acad. Sci. U.S.A.
: VOLUME: 85
: PAGES: 3294-3298
: DATE: May-1988

```

```
US-07-737-736B-5
Query Match          2.6%; Score 30.6; DB 1; Length 1399;
Best Local Similarity 52.8%; Pred. No. 6.5;
Matches 66; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

QY 694 GTCCTTTTTCGGGAGTAGATCCACCTGAACGATAAGCGGAACATTGCTGCTGATG 753
DB 331 GTTCTCCTTGTGTAGTCGGGAGTCCCGGTTGAAGGGCGAGTGAATAGTCCCTTCGCTT 272
QY 754 CAGCCACAGGCGTGATATGACAGGGGTAAACGTCGCGATGGCTTATTATGCAATGCGGGCC 813
DB 271 CATGCTTCGCTGAAGAAGCCTTTTCAGCCTTCACAGGTATGAGTGAAGTGAAGGCC 212
QY 814 AGTGG 818
DB 211 AGTGG 207

RESULT 25
US-09-651-656-26/c
; Sequence 26, Application US/09651656
; Patent No. 6340566
; GENERAL INFORMATION:
; APPLICANT: MCCUTHEN-MALONEY, SANDRA
; APPLICANT: LAWRENCE LIVERMORE NATIONAL LABORATORY
; TITLE OF INVENTION: DETECTION AND QUANTIFICATION OF SINGLE NUCLEOTIDE
; TITLE OF INVENTION: POLYMORPHISMS, DNA SEQUENCE VARIATIONS, DNA MUTATIONS,
; TITLE OF INVENTION: DNA DAMAGE AND DNA MISMATCHES
; FILE REFERENCE: IL-10689
; CURRENT APPLICATION NUMBER: US/09/651,656
; CURRENT FILING DATE: 2000-08-29
; PRIOR APPLICATION NUMBER: 60/192,764
; PRIOR FILING DATE: 2000-03-28
; NUMBER OF SEQ ID NOS: 106
; SOFTWARE: Patent In Ver. 2.1
; LENGTH: 780
; TYPE: DNA
; ORGANISM: Escherichia coli
US-09-651-656-26

Query Match          2.6%; Score 30.4; DB 4; Length 780;
Best Local Similarity 61.2%; Pred. No. 5.4;
Matches 49; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

QY 323 ATCCGCGAGTGGCGGTCATGCAGACATCCCGCCAGGTAACAGCGTCCCTGTGCAC 382
DB 291 AACGCCAGTTCAAGCATCGCTGAGCGGTATTCGCCACCGGTAGATTTCGCCGTGCG 232
QY 383 ATTCTTCTGAAATGACATCAG 402
DB 231 CITATTAACTGACATCGG 212

RESULT 26
US-09-650-855-26/c
; Sequence 26, Application US/09650855
; Patent No. 6365355
; GENERAL INFORMATION:
; APPLICANT: MCCUTHEN-MALONEY, SANDRA
; APPLICANT: LAWRENCE LIVERMORE NATIONAL LABORATORY
; TITLE OF INVENTION: CHIMERIC PROTEINS FOR DETECTION AND QUANTIFICATION OF DNA
; TITLE OF INVENTION: MUTATIONS, DNA SEQUENCE VARIATIONS, DNA DAMAGE AND DNA
; TITLE OF INVENTION: MISMATCHES
; FILE REFERENCE: IL-10284
; CURRENT APPLICATION NUMBER: US/09/650,855
; CURRENT FILING DATE: 2000-08-29
; PRIOR APPLICATION NUMBER: 60/192,764
; PRIOR FILING DATE: 2000-03-28
; NUMBER OF SEQ ID NOS: 106
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 26

US-09-651-656-26/c
; Sequence 26, Application US/09651656
; Patent No. 6340566
; GENERAL INFORMATION:
; APPLICANT: MCCUTHEN-MALONEY, SANDRA
; APPLICANT: LAWRENCE LIVERMORE NATIONAL LABORATORY
; TITLE OF INVENTION: DETECTION AND QUANTIFICATION OF SINGLE NUCLEOTIDE
; TITLE OF INVENTION: POLYMORPHISMS, DNA SEQUENCE VARIATIONS, DNA MUTATIONS,
; TITLE OF INVENTION: DNA DAMAGE AND DNA MISMATCHES
; FILE REFERENCE: IL-10689
; CURRENT APPLICATION NUMBER: US/09/651,656
; CURRENT FILING DATE: 2000-08-29
; PRIOR APPLICATION NUMBER: 60/192,764
; PRIOR FILING DATE: 2000-03-28
; NUMBER OF SEQ ID NOS: 106
; SOFTWARE: Patent In Ver. 2.1
; LENGTH: 780
; TYPE: DNA
; ORGANISM: Escherichia coli
US-09-651-656-26

Query Match          2.6%; Score 30.4; DB 4; Length 780;
Best Local Similarity 61.2%; Pred. No. 5.4;
Matches 49; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

QY 323 ATCCGCGAGTGGCGGTCATGCAGACATCCCGCCAGGTAACAGCGTCCCTGTGCAC 382
DB 291 AACGCCAGTTCAAGCATCGCTGAGCGGTATTCGCCACCGGTAGATTTCGCCGTGCG 232
QY 383 ATTCTTCTGAAATGACATCAG 402
DB 231 CITATTAACTGACATCGG 212

RESULT 27
US-08-682-847-1/c
; Sequence 1, Application US/08682847
; Patent No. 5858989
; GENERAL INFORMATION:
; APPLICANT: BABIUK, LORNE
; APPLICANT: VAN DEN HURK, SYLVIA
; APPLICANT: FITZPATRICK, DAVID
; APPLICANT: ZAMB, TIM
; TITLE OF INVENTION: RECOMBINANT BOVINE HERPESVIRUS TYPE 1
; TITLE OF INVENTION: POLYPEPTIDES AND VACCINES
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 755 PAGE MILL ROAD
; CITY: PALO ALTO
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/682,847
; FILING DATE: 12-JUL-1996
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: PARK, FREDDIE K.
; REGISTRATION NUMBER: 35,636
; REFERENCE/DOCKET NUMBER: 29310-20005.10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 813-5600
; TELEFAX: (415) 494-0792
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3382 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 432..3230
; US-08-682-847-1

Query Match          2.6%; Score 30.4; DB 2; Length 3382;
Best Local Similarity 48.3%; Pred. No. 13;
Matches 85; Conservative 0; Mismatches 91; Indels 0; Gaps 0;

QY 284 GCCTGCCATCATCCGGAACGCTTCGGTCGGGATAAATAATCGCGAGTGGCGCGGTCCA 343
DB 261 GCGCGTGACGCCCGCGCCCGCGGATGCGGAAGCAAAAAAAGTCTGTGAAGTCTGCTAC 202
```

```

QY 344 TGCAGACATCCCGGAGTAAACAGCGTCCCTGTCACATCTTCTGAATGACATCAGG 403
Db 201 CGCCACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 142
QY 404 GATCCCGCGCGTCTCACTGCGGATAACGGCGCGGAGACGACGCTTACGCCA 459
Db 141 GACGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 86

```

RESULT 28

```

US-08-765-907A-4
; Sequence 4, Application US/08765907A
; Patent No. 6352839
; GENERAL INFORMATION:
; APPLICANT: BLANC, Veronique
; APPLICANT: THIBAUT, Denis
; APPLICANT: BAMES-JACQUES, Nathalie
; APPLICANT: BLANCHE, Francis
; APPLICANT: COUZET, Joel
; APPLICANT: BARRIERE, Jean-Claude
; APPLICANT: DEBUSSCHE, Laurent
; APPLICANT: FAMECHON, Alain
; APPLICANT: PARIS, Jean-Marc
; APPLICANT: DUTRUC-ROSSET, Gilles
; TITLE OF INVENTION: Streptogramins And Method For Preparing Same By
; FILE REFERENCE: Streptogramin genes
; CURRENT APPLICATION NUMBER: US/08/765,907A
; CURRENT FILING DATE: 1997-03-20
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 387
; TYPE: DNA
; ORGANISM: Streptomyces pristinaespiralis
US-08-765-907A-4

```

```

Query Match 2.6%; Score 30.2; DB 4; Length 387;
Best Local Similarity 46.1%; Pred. No. 4.2;
Matches 101; Conservative 0; Mismatches 118; Indels 0; Gaps 0;

QY 252 TCAACCGGTACTCAATATCTTCTGCGCGCTGCGCGTCCATCGGATCCCGGAGCGTCCGGT 311
Db 65 TCGACCGGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 124
QY 312 CGGGATAAAAAATCGCGGAGTGGCGGCTGCGCGTCCATGACACATCCCGGAGTAAACAGC 371
Db 125 CAGCGCTGGACATCTGCTGCGCATCGCGGAGTAAAGCGCTCCACCGGTCCGATGA 184
QY 372 GTCCCTGTACATCTTCTGAATGACATCAGGATCCCGCGCTCTCACTGGCGATAACG 431
Db 185 TGCAGCGCGCGGATCGCGGATCGCGGATCGCGGATCGCGGATCGCGGATCGCGGATCG 244
QY 432 GGCAGCGCGGAGACTGACGCTTCAGGACGATACCATACCA 470
Db 245 GCATCGACCGCGCGCTTCTGCGGACCGCTGTACGACACGA 283

```

RESULT 29

```

US-08-765-907A-1
; Sequence 1, Application US/08765907A
; Patent No. 6352839
; GENERAL INFORMATION:
; APPLICANT: BLANC, Veronique
; APPLICANT: THIBAUT, Denis
; APPLICANT: BAMES-JACQUES, Nathalie
; APPLICANT: BLANCHE, Francis
; APPLICANT: COUZET, Joel
; APPLICANT: BARRIERE, Jean-Claude
; APPLICANT: DEBUSSCHE, Laurent
; APPLICANT: FAMECHON, Alain
; APPLICANT: PARIS, Jean-Marc

```

```

; APPLICANT: DUTRUC-ROSSET, Gilles
; TITLE OF INVENTION: Streptogramins And Method For Preparing Same By
; FILE REFERENCE: Mutasynthesis
; CURRENT APPLICATION NUMBER: US/08/765,907A
; CURRENT FILING DATE: 1997-03-20
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 2888
; TYPE: DNA
; ORGANISM: Streptomyces pristinaespiralis
US-08-765-907A-1

Query Match 2.6%; Score 30.2; DB 4; Length 2888;
Best Local Similarity 46.1%; Pred. No. 13;
Matches 101; Conservative 0; Mismatches 118; Indels 0; Gaps 0;

QY 252 TCAACCGGTACTCAATATCTTCTGCGCGCTGCGCGTCCATCGGATCCCGGAGCGTCCGGT 311
Db 1937 TCGACCGGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1996
QY 312 CGGGATAAAAAATCGCGGAGTGGCGGCTGCGCGTCCATCGGATCCCGGAGTAAACAGC 371
Db 1997 CAGCGCTGGACATCTGCTGCGCATCGCGGAGTAAAGCGCTCCACCGGTCCGATGA 2056
QY 372 GTCCCTGTACATCTTCTGAATGACATCAGGATCCCGCGCTCTCACTGGCGATAACG 431
Db 2057 TGCAGCGCGCGGATCGCGGATCGCGGATCGCGGATCGCGGATCGCGGATCGCGGAT 2116
QY 432 GGCAGCGCGGAGACTGACGCTTCAGGACGATACCATACCA 470
Db 2117 GCATCGACCGCGCGCTTCTGCGGACCGCTGTACGACACGA 2155

```

RESULT 30

```

US-09-041-886-24
; Sequence 24, Application US/09041886
; Patent No. 6235872
; GENERAL INFORMATION:
; APPLICANT: Bredesen, Dale E.
; APPLICANT: Rabizadeh, Sharroz
; TITLE OF INVENTION: Proapoptotic peptides, dependence
; TITLE OF INVENTION: Polypeptides and Methods of Use
; NUMBER OF SEQUENCES: 72
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell & Flores LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: United States
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/041,886
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-LJ 2626
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4608 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single

```


Db 260 TNSNTNCCNWSNMGNCAYMGNTNTTYTNNACNTAYMGNCAITGTMGNAAYTTWSNA 319
QY 679 GCAGCATAAATTCATGCTCTTTTCGGGACGTAAGCATCCCACTGAACGATAAGCGGAA 738
Db 320 THYNTNTGARCCNWSNGGNTGWSNAARGAYACNTTYTNTYNTNGCNATHAARWSNC 379
QY 739 CATTGCTGCTGATGACGCCAGCGCTGGATATGACGGGTAAACGGTCCGATGCTTCAT 798
Db 380 ARCCNGNCAIYNTGARMGNGMGCNCGNATHMGNWSNACNTGGGNGMNGTNGGNGNT 439
QY 799 TATGCAATGGGCGGAGTGAACACCGGTGGTAATAACCGCTTACCGGTGCTCCTGACACCTT 858
Db 440 GGCNMGNGMGNCAIYNTNAARYTNGTNTTYTNTNGGNGTNGCNGGWSNCCNC 499
QY 859 CCGC 862
Db 500 CNGC 503

RESULT 33
US-08-572-951-3/c
; Sequence 3, Application US/08572951
; Patent No. 5824790
; GENERAL INFORMATION:
; APPLICANT: KEELING, PETER L.
; APPLICANT: KNIGHT, MARY E.
; APPLICANT: GUAN, HANPING
; TITLE OF INVENTION: MODIFICATION OF STARCH
; TITLE OF INVENTION: SYNTHESIS IN PLANTS
; NUMBER OF SEQUENCES: 41
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN DABBY & CUSHMAN
; ADDRESSEE: Intellectual Property Group of
; ADDRESSEE: Pillsbury Madison & Sutro LLP
; STREET: 1100 New York Avenue, N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3918
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/572,951
; FILING DATE: 15-DEC-1995
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/346,602
; FILING DATE: 29-NOV-1994
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/263,921
; FILING DATE: 21-JUN-1994
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Paul N. Kokulis
; REGISTRATION NUMBER: 16,773
; REFERENCE/DOCKET NUMBER: 222957/1.02.15C
; TELEPHONE: (202) 861-3000
; TELEFAX: (202) 822-0944
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2380 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-572-951-3

Query Match 2.5%; Score 29.2; DB 1; Length 2380;

Best Local Similarity 28.9%; Pred. No. 26;
Matches 110; Conservative 42; Mismatches 222; Indels 6; Gaps 1;
QY 398 ATCAGGATCCCGCGCTCTCACTGGCGATAAGCGGCACCGCGAGACTGACGCTTCAGC 457
Db 1926 RTCNCKNARNCCNCCNACNCRGTGNACNCGNACNGTNCRTANGCCATNGCRTANAR 1867
QY 458 CAGTACATACCAAAAGCTTCATTTCCGAAGGATGACCACTGCAATGCCGTA 517
Db 1866 YTGRTTNARCCRCANGYTCRAANCKNSWGGCATNARNARDATRTCNCGNCCNGT 1807
QY 518 GACCGGTAAAGCTGGGAAAGGCACCTGCTTAACACATCTCCGCTCATCCAGGTG 577
Db 1806 DATNCKRTGMCNARNGNACNSWRAANCCNACCCANGCNCKNACYTTRTCNSWRTGTC 1747
QY 578 TTCTGTCTGTGACGACGAGCTGCTGTTATTTTCACGCCCGCGCCACCGAGGCCA 637
Db 1746 NSWYTCRAANCKNCKNARCATRTCTCNARRTCNCGNCKNCCNCCNARCATNACNAR 1687
QY 638 GCGAAATGATTTCCCTTCCATCTTCAGCTGATACATACAGGAGCATATAATTCATGTC 697
Db 1686 YTGNACTCTGTGNCNGCDAT-----CCARTGDTNGCRTNCGCDATDATRTCNACNCC 1633
QY 698 TTTTTCGGGACGTAGCATCCCACTGAAAGCGGAGCAATGTCTGCTGATGAGCAGC 757
Db 1632 YTTTGRTGTCNARCKNCCDTRAANCCDTRAANNGNACRTCTCTCNCKNACYTGNAR 1573
QY 758 CCAGGCTGGATATGACGGG 777
Db 1572 NCCNARYTGNCYTGNAAG 1553

RESULT 34
US-08-832-883-48/c
; Sequence 48, Application US/08832883
; Patent No. 5807681
; GENERAL INFORMATION:
; APPLICANT: Giordano, Antonio
; APPLICANT: Baldi, Alphonso
; TITLE OF INVENTION: METHODS FOR THE DIAGNOSIS AND PROGNOSIS
; TITLE OF INVENTION: OF CANCER
; NUMBER OF SEQUENCES: 115
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEIDEL, GONDA, LAVORNA & MONACO, P.C.
; STREET: Suite 1800 Two Penn Center Plaza
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/832,883
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Monaco, Daniel A.
; REGISTRATION NUMBER: 30,480
; REFERENCE/DOCKET NUMBER: 8321-13 US1
; TELEPHONE: (215) 568-8383
; TELEFAX: (215) 568-5549
; INFORMATION FOR SEQ ID NO: 48:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3865 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-832-883-48

Query Match 2.5%; Score 29.2; DB 1; Length 3865;
Best Local Similarity 51.5%; Pred. No. 34;
Matches 67; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

QY 394 TGACATCAGGATCCCGCCGCTCTCACTGGCGATAACGGCAGCGCCGGAGACTGACGCTT 453
DB 1176 TGAATCAGGTGATTCGCGCTCCCTCAGCCTCATAAAGTCTGGGATTACAGGGTGAGC 1117

QY 454 CAGCCAGTACCATACCAAAACGCTTCATTTCCGAAGGCATGACCAACACACTGGCAATCC 513
DB 1116 CAACCCGCTGGCTCAGATGCTTCATTTTAAATAGTCAATCCCAAACTGTAAAGTGA 1057

QY 514 GGTAGACCGG 523
DB 1056 GCTAGACCAG 1047

RESULT 35
US-08-832-877-48/c
; Sequence 48, Application US/08832877
; Patent No. 5840506
; GENERAL INFORMATION:
; APPLICANT: Giordano, Antonio
; TITLE OF INVENTION: METHODS FOR THE DIAGNOSIS AND PROGNOSIS OF
; TITLE OF INVENTION: CANCER
; NUMBER OF SEQUENCES: 116
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEIDEL, GONDA, LAVORGNA & MONACO, P.C.
; STREET: Suite 1800 Two Penn Center Plaza
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/832,877
; FILING DATE:
; CLASSIFICATION: 436
; ATTORNEY/AGENT INFORMATION:
; NAME: Monaco, Daniel A
; REGISTRATION NUMBER: 30,480
; REFERENCE/DOCKET NUMBER: 8321-13 US2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-8383
; TELEFAX: (215) 568-5549
; INFORMATION FOR SEQ ID NO: 48:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3865 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-832-877-48

Query Match 2.5%; Score 29.2; DB 2; Length 3865;
Best Local Similarity 51.5%; Pred. No. 34;
Matches 67; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

QY 394 TGACATCAGGATCCCGCCGCTCTCACTGGCGATAACGGCAGCGCCGGAGACTGACGCTT 453
DB 1176 TGAATCAGGTGATTCGCGCTCCCTCAGCCTCATAAAGTCTGGGATTACAGGGTGAGC 1117

QY 454 CAGCCAGTACCATACCAAAACGCTTCATTTCCGAAGGCATGACCAACACACTGGCAATCC 513
DB 1116 CAACCCGCTGGCTCAGATGCTTCATTTTAAATAGTCAATCCCAAACTGTAAAGTGA 1057

QY 514 GGTAGACCGG 523
DB 1056 GCTAGACCAG 1047

Db 1056 GCTAGACCAG 1047

RESULT 36
US-09-293-505-8
; Sequence 8, Application US/09293505
; Patent No. 6348575
; GENERAL INFORMATION:
; APPLICANT: de Sauvage, Frederic
; APPLICANT: Carpenter, David A.
; TITLE OF INVENTION: Patched-2
; FILE REFERENCE: P1405R1
; CURRENT APPLICATION NUMBER: US/09/293,505
; CURRENT FILING DATE: 1999-04-15
; EARLIER APPLICATION NUMBER: US 60/081,884
; EARLIER FILING DATE: 1998-04-15
; NUMBER OF SEQ ID NOS: 32
; SEQ ID NO 8
; LENGTH: 4004
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-293-505-8

Query Match 2.5%; Score 29.2; DB 4; Length 4004;
Best Local Similarity 62.2%; Pred. No. 35;
Matches 46; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

QY 411 CCGTCTCACTGGCGATAACGGCAGCGCCGGAGACTGACGCTTCAGCCAGTACCA 470
DB 3577 CAGTGGCTGTGGTACGACGGGCCCCCGGGGAGGACCAATCAGCTGATTCAGTATTC 3636

QY 471 AACGCTTCATTTTC 484
DB 3637 AACACATATGTTTC 3650

RESULT 37
5405943-1/c
; Patent No. 5405943
; APPLICANT: COMINGS, DAVID E.
; TITLE OF INVENTION: TOURETTE SYNDROM, AUTISM AND ASSOCIATED
; BEHAVIORS
; NUMBER OF SEQUENCES: 5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/562,596
; FILING DATE: 03-AUG-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 125,577
; FILING DATE: 25-NOV-1987
; APPLICATION NUMBER: 271,653
; FILING DATE: 16-NOV-1988
; APPLICATION NUMBER: 410,831
; FILING DATE: 22-SEP-1989
; SEQ ID NO:1:
; LENGTH: 748
5405943-1

Query Match 2.5%; Score 29; DB 6; Length 748;
Best Local Similarity 51.1%; Pred. No. 15;
Matches 68; Conservative 0; Mismatches 65; Indels 0; Gaps 0;

QY 1013 ATTACTGTCACTGGCCACTATGCGAGATGACGAGATGACGAGTAAATCCCGGATAATCCG 1072
DB 345 ATCTCTGACATGCGGATTTCTGAAAGATCTCTGAAACAGATCCCAACTCCAGAGATTTG 286

QY 1073 TCGAAGTCTGAGGATGGAAGAGTGAAGGCTGTCTCTGAAGGAATAAAGTGACATC 1132
DB 285 CTAAACACAGATTCATTAAGCTTGTATGATTAATAAAAGATGTTTCATCATGATTTT 226

QY 1133 ATGCCCTCTTTTT 1145
DB 225 ATTCTCTTTTGT 213

```

RESULT 38
US-08-858-207A-78/c
; Sequence 78, Application US/08858207A
; Patent No. 6348328
; GENERAL INFORMATION:
; APPLICANT: Black, Michael
; APPLICANT: Hodgson, John
; APPLICANT: Knowles, David
; APPLICANT: Nicholas, Richard
; APPLICANT: Stodola, Robert
; TITLE OF INVENTION: No. 6348328el Compounds
; NUMBER OF SEQUENCES: 552
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Smithline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/858,207A
; FILING DATE: 09-MAY-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/017670
; FILING DATE: 14-MAY-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Glimmi, Edward R
; REGISTRATION NUMBER: 38,891
; REFERENCE/DOCKET NUMBER: P50475
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-4478
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 78:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1426 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-858-207A-78

Query Match 2.5%; Score 29; DB 4; Length 1426;
Best Local Similarity 55.4%; Pred. No. 22;
Matches 56; Conservative 0; Mismatches 45; Indels 0; Gaps 0;

QY 173 GAGGATCCGCGTTATCTGTTGCAATTTCCCTTAGCTGACTAGCCAGACAGACACATGAT 232
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 905 GATGATATGCCCTTGTGCTTCTTCATACGCGCGGCGGAGGTGAGGACGACAGATTCAT 846

QY 233 CTGTGCGGTTCCTTTAATATCAAAACCGGTACTCAATATCTT 273
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 845 CAGGCGAAAGGTATTGACCTCAAAACATCTGGTGAATATCTT 805

RESULT 39
US-08-961-527-50
; Sequence 50, Application US/08961527
; Patent No. 6420135
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 391
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: MD
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44Mb storage
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 8.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/453,702B
; FILING DATE: 03-Dec-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/110,955

```

```

; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44Mb storage
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,527
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 50:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 538 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; US-08-961-527-50

Query Match 2.5%; Score 29; DB 4; Length 538;
Best Local Similarity 55.4%; Pred. No. 48;
Matches 56; Conservative 0; Mismatches 45; Indels 0; Gaps 0;

QY 173 GAGGATCCGCGTTATCTGTTGCAATTTCCCTTAGCTGACTAGCCAGACAGACACATGAT 232
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1327 GATGATATGCCCTTGTGCTTCTTCATACGCGCGGCGGAGGTGAGGACGACAGATTCAT 1386

QY 233 CTGTGCGGTTCCTTTAATATCAAAACCGGTACTCAATATCTT 273
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1387 CAGGCGAAAGGTATTGACCTCAAAACATCTGGTGAATATCTT 1427

RESULT 40
US-09-453-702B-39/c
; Sequence 39, Application US/09453702B
; Patent No. 6365723
; GENERAL INFORMATION:
; APPLICANT: Blattner, Frederick R.
; Burland, Valerie
; Perna, Nicole T.
; Plunkett, Guy
; Welch, Rod
; TITLE OF INVENTION: No. 6365723el Sequences of E. coli O157
; NUMBER OF SEQUENCES: 265
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Quarles & Brady
; STREET: 1 South Pinckney Street
; CITY: Madison
; STATE: WI
; COUNTRY: USA
; ZIP: 53701-2113
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44Mb storage
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 8.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/453,702B
; FILING DATE: 03-Dec-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/110,955

```

;; FILING DATE: 04-DEC-1998
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Seay, Nicholas J.
;; REGISTRATION NUMBER: 27386
;; REFERENCE/DOCKET NUMBER: 960296.95017
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (608) 251-5000
;; TELEFAX: (608) 251-9166
;; INFORMATION FOR SEQ ID NO: 39:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 25165
;; TYPE: nucleic acid
;; STRANDEDNESS: double
;; TOPOLOGY: linear
;; MOLECULE TYPE: DNA (genomic)
;; SEQUENCE DESCRIPTION: SEQ ID NO: 39:
US-09-453-702B-39

Query Match 2.5%; Score 29; DB 4; Length 25165;
Best Local Similarity 58.8%; Pred. No. 1.2e+02;
Matches 50; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

QY 380 CACATCTCTCGAATGACATCAGGATCCCGCGCTCTCACTGCGGATACGGGCGAGCC 439
DB 4215 CACATTATTCAGGTCGCGGTGACGATATCGCGCGCTGAGCGCGCAGTACCGTGGCGCT 4156

QY 440 GGAGACTGACGTCGAGCCATGACC 464
DB 4155 GATGATTTGCGCGCTGAGCATGTTCC 4131

RESULT 41
US-08-578-634C-2
; Sequence 2, Application US/08578634C
; Patent No. 6025163
; GENERAL INFORMATION:
; APPLICANT: Vladimir Shamanin
; APPLICANT: Ethel Michele De Villiers
; APPLICANT: Zur Hausen
; APPLICANT: Irene Leigh
; APPLICANT: Harald Zur Hausen
; TITLE OF INVENTION: DNA CODING FOR A PEPTIDE OF A PAPILLOMA
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of The Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/578,634C
; FILING DATE: 26-JAN-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Abrams, Samuel B.
; REGISTRATION NUMBER: 30,605
; REFERENCE/DOCKET NUMBER: 8484-007
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 493-4935
; TELEFAX: (650) 493-5556
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:

;; LENGTH: 668 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: double
;; TOPOLOGY: linear
;; MOLECULE TYPE: DNA
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: 1 .. 666
US-08-578-634C-2

Query Match 2.4%; Score 28.6; DB 3; Length 668;
Best Local Similarity 57.1%; Pred. No. 19;
Matches 52; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY 910 ACAGATAATTTCAGGGAACACGTTCTTACGGGTGATGTAGTGTCTCTGACAA 969
DB 240 ACTGGTAAATACTGTAAATAGAAGATGGAGATATGTTGGATATAGGTTTGGAAATATAA 299

QY 970 TAGTGAAGCGGTGACAGCATATCAGACGGCT 1000
DB 300 TAATRAAACCCGTGTCAGCAATAAGTCAGAT 330

RESULT 42
US-09-430-010-2
; Sequence 2, Application US/09430010
; Patent No. 6395512
; GENERAL INFORMATION:
; APPLICANT: Vladimir Shamanin
; APPLICANT: Ethel Michele De Villiers
; APPLICANT: Zur Hausen
; APPLICANT: Irene Leigh
; APPLICANT: Harald Zur Hausen
; TITLE OF INVENTION: DNA CODING FOR A PEPTIDE OF A PAPILLOMA
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of The Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/430,010
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/578,634
; FILING DATE: 26-JAN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Abrams, Samuel B.
; REGISTRATION NUMBER: 30,605
; REFERENCE/DOCKET NUMBER: 8484-007
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 493-4935
; TELEFAX: (650) 493-5556
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 668 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1 .. 666

US-09-430-010-2

Query Match 2.4%; Score 28.6; DB 4; Length 668;
Best Local Similarity 57.1%; Pred. No. 19;
Matches 52; Conservative 0; Mismatches 39; Indels 0; Gaps 0;
QY 910 ACAGATAAATCAGGAAACGTTCTGCTTACGGGTGATGAGTGTCTTTTGTCTGACAA 969
DB 240 ACTGGTAATACGTAAATGAGATGAGATATGGTGGATATAGTTTGGAAATATATA 299
QY 970 TAGTGAAGCGGTGACAGCATATCAGACGGCT 1000
DB 300 TAATAAAACCTGTCAGCAAAATAGTCAGAT 330

RESULT 43

US-09-575-574-5
Sequence 5, Application US/09575574
Patent No. 6376750
GENERAL INFORMATION:
APPLICANT: Yu, Su-May
APPLICANT: Chao, Yu-Chan
TITLE OF INVENTION: PLANT SEEDLING AND EMBRYO PROMOTER
FILE REFERENCE: 08919-047001
CURRENT APPLICATION NUMBER: US/09/575,574
CURRENT FILING DATE: 2000-05-22
NUMBER OF SEQ ID NOS: 10
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 5
LENGTH: 1730
TYPE: DNA
ORGANISM: Oryza sativa
US-09-575-574-5

Query Match 2.4%; Score 28.6; DB 4; Length 1730;
Best Local Similarity 55.6%; Pred. No. 34;
Matches 55; Conservative 0; Mismatches 44; Indels 0; Gaps 0;
QY 876 ATCATGGGTGATGATGAGTGTCTTTTGTCTGACAAATAGTG 974
DB 343 ATCATGATTAATAGGCTCAATAGATTTGTCTCGGGAATTAGTCGAGATTATGATGA 402
QY 936 GTCTTACGGGTGATGATGAGTGTCTTTTGTCTGACAAATAGTG 974
DB 403 GTTTTATGATGATCTACGTTTAAATATTAATATAGTG 441

RESULT 44

US-08-608-241-5
Sequence 5, Application US/08608241
Patent No. 5747328
GENERAL INFORMATION:
APPLICANT: Donohue, Timothy J
APPLICANT: Barber, Robert D
APPLICANT: Witthuhn, Vernon
TITLE OF INVENTION: MICROBIAL SYSTEM FOR FORMALDEHYDE
TITLE OF INVENTION: SENSING AND REMEDIATION
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Quarles & Brady
STREET: 1 South Pinckney Street
CITY: Madison
STATE: WI
COUNTRY: US
ZIP: 53703
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/608,241
FILING DATE:

CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Seay, Nicholas J
REGISTRATION NUMBER: 27,386
REFERENCE/DOCKET NUMBER: 960296.93511
TELEPHONE: 608-251-5000
TELEFAX: 608-251-9166
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 2563 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORGANISM: Rhodobacter sphaeroides
STRAIN: 2.4.1
IMMEDIATE SOURCE:
CLONE: pUI8017
FEATURE:
NAME/KEY: CDS
LOCATION: 2113..2400
OTHER INFORMATION: /product= "Partial Sensor Coding
OTHER INFORMATION: Sequence"
US-08-608-241-5

Query Match 2.4%; Score 28.6; DB 1; Length 2563;
Best Local Similarity 47.1%; Pred. No. 43;
Matches 88; Conservative 0; Mismatches 99; Indels 0; Gaps 0;
QY 226 CAATGATCTGTGCCGTCTCTGTTAATATCAACCGGTACTCAATATCTTCTGCGCTGG 285
DB 849 CGAGCAGCTGACGCGGAATGCATGTATGTGAACACGACCATGACCGGGTCCGAT 908
QY 286 CTGCCATCATCCGAGAGCGTTCCGGTCCGGATATAAATCGCCAGTGCCTGCGCATG 345
DB 909 CTACCCCGGGGAGGAGCGATGAGGCGGCGCGGGGTGACCTGCTCCCGTGGGG 968
QY 346 CAGACACATCCCCACCGGTAAACAGCGTCCCTGTACATTTCTTCTGAATGACATCAGGA 405
DB 969 AAGGACCATGCGCCACGCGGAGCGGAGGACATGAGCCITCTCTGATCGATCCCGCA 1028
QY 406 TCCCGCC 412
DB 1029 TCCGCC 1035

RESULT 45

US-08-922-182-5
Sequence 5, Application US/08922182
Patent No. 5834300
GENERAL INFORMATION:
APPLICANT: Donohue, Timothy J
APPLICANT: Barber, Robert D
APPLICANT: Witthuhn, Vernon
TITLE OF INVENTION: MICROBIAL SYSTEM FOR FORMALDEHYDE
TITLE OF INVENTION: SENSING AND REMEDIATION
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Quarles & Brady
STREET: 1 South Pinckney Street
CITY: Madison
STATE: WI
COUNTRY: US
ZIP: 53703
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/922,182

```

RESULT 46
US-08-919-953-5
; Sequence 5, Application US/08919953
; Patent No. 5837481
; GENERAL INFORMATION:
; APPLICANT: Donohue, Timothy J
; APPLICANT: Barber, Robert D
; APPLICANT: Witthuhn, Vernon
; TITLE OF INVENTION: MICROBIAL SYSTEM FOR FORMALDEHYDE
; TITLE OF INVENTION: SENSING AND REMEDIATION
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Quarles & Brady
; STREET: 1 South Pinckney Street
; CITY: Madison
; STATE: WI
; COUNTRY: US
; ZIP: 53703
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

```

```

RESULT 47
US-09-575-574-3
; Sequence 3, Application US/09575574
; Patent No. 6376750
; GENERAL INFORMATION:
; APPLICANT: Yu, Su-May
; APPLICANT: Chao, Yu-Chan
; TITLE OF INVENTION: PLANT SEEDLING AND EMBRYO PROMOTER
; FILE REFERENCE: 08919-047001
; CURRENT APPLICATION NUMBER: US/09/575,574
; CURRENT FILING DATE: 2000-05-22
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 3572
; TYPE: DNA
; ORGANISM: Oryza sativa

```

US-08-279-058B-13/c
Sequence 13, Application US/08279058B
Patent No. 5668004
GENERAL INFORMATION:
APPLICANT: Michael E. O'Donnell
TITLE OF INVENTION: DNA POLYMERASE III
TITLE OF INVENTION: HOLOENZYME
NUMBER OF SEQUENCES: 60
CORRESPONDENCE ADDRESS:
ADDRESSEE: Yahwak & Associates
STREET: 25 Skytop Drive
CITY: Trumbull
STATE: Connecticut
COUNTRY: USA
ZIP: 06611
COMPUTER READABLE FORM:

Db 574 GCATGGCGCAAGTAATGCATC 553

Search completed: June 19, 2003, 03:36:30
Job time : 73.8861 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 18, 2003, 04:53:09 ; Search time 298.567 Seconds
(without alignments)
5804.490 Million cell updates/sec

Title: US-09-674-277-2
Perfect score: 1181
Sequence: 1 ctccagagatgaaaaa.....ttttactttttctgtcag 1181

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1042519 seqs, 733713590 residues

Total number of hits satisfying chosen parameters: 2085038

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 150 summaries

Database : Published Applications, NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
- 12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
- 13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
- 14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
c 1	49	4.1	1135	10	US-09-974-300-1695
c 2	48.6	4.1	783	10	US-09-974-300-6118
c 3	46.2	3.9	783	10	US-09-974-300-6111
c 4	41.6	3.5	1227	10	US-09-974-300-1771
c 5	40.6	3.4	129908	9	US-10-270-875-1
c 6	40.6	3.4	129908	9	US-10-270-878-1
c 7	40.6	3.4	129908	9	US-10-270-786-1
c 8	40.6	3.4	129908	9	US-10-270-710-1
c 9	40.6	3.4	129908	9	US-10-270-859-1
c 10	37	3.1	1108	10	US-09-974-300-1730
c 11	36.6	3.1	598	10	US-09-825-561A-7
c 12	36.2	3.1	598	10	US-09-864-761-14748
c 13	36.2	3.1	741	10	US-09-825-561A-66
c 14	36	3.0	1427	10	US-09-974-300-757
c 15	35.4	3.0	480	9	US-09-918-995-8152
c 16	35	3.0	1135	10	US-09-974-300-1702
c 17	34.8	2.9	513	9	US-10-123-155-536
c 18	34.8	2.9	802	9	US-10-184-644-312
c 19	34.8	2.9	802	9	US-10-184-634-312

20	34.2	2.9	74962	9	US-10-274-974-3	Sequence 3, Appl1
c 21	33.8	2.9	527	9	US-10-184-644-522	Sequence 522, App
c 22	33.8	2.9	527	9	US-10-184-634-522	Sequence 522, App
c 23	33.8	2.9	1686	12	US-10-044-090-362	Sequence 362, App
c 24	33.8	2.9	1712	10	US-09-880-107-3347	Sequence 3347, App
c 25	33.2	2.8	716	9	US-10-123-155-96	Sequence 96, Appl1
c 26	33.2	2.8	777	9	US-10-184-644-348	Sequence 348, App
c 27	33.2	2.8	777	9	US-10-184-634-348	Sequence 348, App
c 28	33	2.8	690	9	US-10-123-155-306	Sequence 306, App
c 29	32.8	2.8	429	9	US-09-918-995-5997	Sequence 5997, App
c 30	32.6	2.8	617	9	US-10-198-846-9370	Sequence 9370, App
c 31	32.4	2.7	811	9	US-10-184-644-414	Sequence 414, App
c 32	32.4	2.7	811	9	US-10-184-634-414	Sequence 414, App
c 33	32.4	2.7	11649	9	US-09-764-891-6047	Sequence 6047, App
c 34	32.4	2.7	11649	9	US-10-091-438-272	Sequence 272, App
c 35	32.2	2.7	468	9	US-10-123-155-90	Sequence 90, Appl1
c 36	32.2	2.7	1254	9	US-09-738-626-451	Sequence 451, App
c 37	32.2	2.7	1689	10	US-09-917-800A-1734	Sequence 1734, App
c 38	32.2	2.7	3309400	9	US-09-738-626-1	Sequence 1, Appl1
c 39	32	2.7	1311	10	US-09-815-242-4609	Sequence 4609, App
c 40	32	2.7	1563	10	US-09-815-242-8107	Sequence 8107, App
c 41	31.8	2.7	653	9	US-10-184-644-402	Sequence 402, App
c 42	31.8	2.7	653	9	US-10-184-634-402	Sequence 402, App
c 43	31.8	2.7	1049	9	US-10-123-155-358	Sequence 358, App
c 44	31.8	2.7	1146	9	US-09-891-641-35	Sequence 35, Appl1
c 45	31.8	2.7	3153	9	US-09-759-130B-343	Sequence 343, App
c 46	31.8	2.7	3153	9	US-10-189-123-73	Sequence 73, Appl1
c 47	31.6	2.7	671	9	US-10-184-644-346	Sequence 346, App
c 48	31.6	2.7	671	9	US-10-184-634-346	Sequence 346, App
c 49	31.6	2.7	2199	10	US-09-381-624A-5	Sequence 5, Appl1
c 50	31.6	2.7	2817	10	US-09-381-624A-4	Sequence 4, Appl1
c 51	31.4	2.7	776	10	US-09-902-331-1	Sequence 1, Appl1
c 52	31.2	2.6	592	9	US-10-198-846-9001	Sequence 9001, App
c 53	31.2	2.6	2176	9	US-09-843-250-8	Sequence 8, Appl1
c 54	31.2	2.6	2774	10	US-09-974-298-23	Sequence 23, Appl1
c 55	31.2	2.6	3664	10	US-09-880-107-3416	Sequence 3416, App
c 56	31.2	2.6	8802	10	US-09-986-900-1	GENERAL INFORMA
c 57	31.2	2.6	14462	9	US-09-843-250-9	Sequence 9, Appl1
c 58	31.2	2.6	32192	9	US-10-092-154-1416	Sequence 1416, App
c 59	31.2	2.6	32192	9	US-09-764-891-7945	Sequence 7945, App
c 60	31.2	2.6	32192	9	US-09-764-891-8220	Sequence 8220, App
c 61	31.2	2.6	32192	9	US-09-764-891-8319	Sequence 8319, App
c 62	31.2	2.6	32192	10	US-09-764-847-1416	Sequence 1416, App
c 63	31.2	2.6	34337	10	US-09-741-149-3	Sequence 3, Appl1
c 64	31	2.6	734	9	US-10-184-644-458	Sequence 458, App
c 65	31	2.6	734	9	US-10-184-634-458	Sequence 458, App
c 66	31	2.6	1440	9	US-09-891-641-47	Sequence 47, Appl1
c 67	31	2.6	2652	9	US-10-195-970-4	Sequence 4, Appl1
c 68	31	2.6	2652	9	US-10-195-970-5	Sequence 5, Appl1
c 69	31	2.6	22299	9	US-10-091-572-469	Sequence 469, App
c 70	31	2.6	22299	9	US-09-764-891-6605	Sequence 6605, App
c 71	30.8	2.6	360	10	US-09-764-877-587	Sequence 587, App
c 72	30.8	2.6	491	9	US-09-918-995-15312	Sequence 15312, A
c 73	30.8	2.6	4385	9	US-10-050-704-77	Sequence 77, Appl1
c 74	30.8	2.6	4385	9	US-10-050-704-23	Sequence 23, Appl1
c 75	30.8	2.6	4386	9	US-10-050-704-78	Sequence 78, Appl1
c 76	30.8	2.6	5511	9	US-09-974-298-13	Sequence 13, Appl1
c 77	30.8	2.6	5511	12	US-10-044-090-285	Sequence 285, App
c 78	30.6	2.6	300	9	US-10-184-644-492	Sequence 492, App
c 79	30.6	2.6	300	9	US-10-184-634-492	Sequence 492, App
c 80	30.6	2.6	451	9	US-10-123-155-126	Sequence 126, App
c 81	30.6	2.6	468	9	US-10-123-155-90	Sequence 90, Appl1
c 82	30.6	2.6	651	10	US-09-811-284-26	Sequence 26, Appl1
c 83	30.6	2.6	705	10	US-09-974-300-2795	Sequence 2795, App
c 84	30.6	2.6	1114	10	US-09-764-864-506	Sequence 506, App
c 85	30.6	2.6	1548	9	US-10-037-270-965	Sequence 965, App
c 86	30.6	2.6	1802	9	US-09-892-877-70	Sequence 70, Appl1
c 87	30.6	2.6	1802	9	US-09-948-783-70	Sequence 70, Appl1
c 88	30.6	2.6	5957	10	US-10-125-540-627	Sequence 627, App
c 89	30.6	2.6	5957	10	US-09-764-870-627	Sequence 627, App
c 90	30.4	2.6	175	9	US-10-025-380-873	Sequence 873, App
c 91	30.4	2.6	175	10	US-09-922-217-873	Sequence 873, App
c 92	30.4	2.6	175	10	US-09-833-263-873	Sequence 873, App

Db 631 TTCCCAAGGGGCGATTAATAACAGTTACTTCTGCTTCAACAACTTCGGGATGCCACCA 572
QY 414 GTCTACTGGCGATACGGGACGGCGGAGACTGACGCTTACGCCAGTACCATACCAAC 473
Db 571 ATTTTGTGCGCATCAGAGTACCGCGCATGCCATCGCTTCAAGCGCCACTAAACCAAG 512
QY 474 GCTTCAATTTCCGAAGCGATGACCAACACACTGGCAATCCGGTAGACCGGTACGCTGG 533
Db 511 CTTTCTTTTTCGCTTAAACACAGCAATTAATCGCTCATAGACAACAATTCAGCAACATGC 452
QY 534 AAAAGGACCTGCCATCACTCCGCTCATTCCTCCAGGCTGTTCTGCTGCTGACGC 593
Db 451 TTTTGTGCTGCTAGCATAGGACATGTTTGTCACGCCGTTTCTTCCACAAACTCCCTC 392
QY 594 AGACGTG 600
Db 391 GCAATTG 385

RESULT 3

US-09-974-300-6111/c
; Sequence 6111, Application US/09974300
; Patent No. US20020146721A1
; GENERAL INFORMATION:
; APPLICANT: Berka, Randy M.
; APPLICANT: Clausen, Ib Groth
; TITLE OF INVENTION: Methods For Monitoring Multiple Gene
; TITLE OF INVENTION: Expression
; FILE REFERENCE: 10085.500-US
; CURRENT APPLICATION NUMBER: US/09/974,300
; PRIOR FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 09/680,598
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/279,526
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 8481
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6111
; LENGTH: 783
; TYPE: DNA
; ORGANISM: Bacillus clausii
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(783)
; OTHER INFORMATION: n = A,T,C or G
US-09-974-300-6111

Query Match 3.9%; Score 46.2; DB 10; Length 783;
Best Local Similarity 59.5%; Pred. No. 4.5e-05;
Matches 78; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

QY 354 TCCCCACGGGTAAACAGCGTCCCTGTACATTTCTTGAATGACATCAGGCGATCCGCCCC 413
Db 633 TTCCCAAGGGGCGATTAATAACAGTTACTTCTGCTTCAACAACTTCGGGATGCCACCA 574
QY 414 GTCTACTGGCGATACGGGACGGCGGAGACTGACGCTTACGCCAGTACCATACCAAC 473
Db 573 ATTTTGTGCGCATCAGAGTACCGCGCATGCCATCGCTTCAAGCGCCACTAAACCAAG 514
QY 474 GCTTCAATTTTC 484
Db 513 CTTTCTTTTC 503

RESULT 4

US-09-974-300-1771/c
; Sequence 1771, Application US/09974300
; Patent No. US20020146721A1
; GENERAL INFORMATION:
; APPLICANT: Berka, Randy M.
; APPLICANT: Clausen, Ib Groth
; TITLE OF INVENTION: Methods For Monitoring Multiple Gene
; TITLE OF INVENTION: Expression

; FILE REFERENCE: 10085.500-US
; CURRENT APPLICATION NUMBER: US/09/974,300
; PRIOR FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 09/680,598
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/279,526
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 8481
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1771
; LENGTH: 1227
; TYPE: DNA
; ORGANISM: Bacillus licheniformis
US-09-974-300-1771

Query Match 3.5%; Score 41.6; DB 10; Length 1227;
Best Local Similarity 62.5%; Pred. No. 0.0024;
Matches 65; Conservative 0; Mismatches 39; Indels 0; Gaps 0;
QY 352 CATCCCCCAGGGTAACAGGTCCTCCGTGCACATCTTCTGAATGACATCAGGATCCCGC 411
Db 1072 CATCTCTTTAGGGGTACCATACCGGTGCACATTTATGTTCAAGATTCGGGAAGTCCCC 1013
QY 412 CGTCTCTACTGGCGATAACGGGACCGCGGAGACTGACGCTTCA 455
Db 1012 CTGCATCGCTGACGATAATAGGAAGTCCGGCTAATTGCGCTTCA 969

RESULT 5

US-10-270-875-1/c
; Sequence 1, Application US/10270875
; Publication No. US20030082741A1
; GENERAL INFORMATION:
; APPLICANT: Sigrður Hjorleifsdóttir
; APPLICANT: Gudmundur O. Hreggvidsson
; APPLICANT: Olafur H. Fridjonsson
; APPLICANT: Arthor Aevarsson
; APPLICANT: Jakob K. Kristjansson
; TITLE OF INVENTION: Bacteriophage RM378 of a Thermophilic
; TITLE OF INVENTION: Host Organism
; FILE REFERENCE: 2739.1001-001
; CURRENT APPLICATION NUMBER: US/10/270,875
; CURRENT FILING DATE: 2002-10-11
; PRIOR APPLICATION NUMBER: US/09/585,858
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: 60/137,120
; PRIOR FILING DATE: 1999-06-02
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 129908
; TYPE: DNA
; ORGANISM: Bacteriophage RM378
US-10-270-875-1

Query Match 3.4%; Score 40.6; DB 9; Length 129908;
Best Local Similarity 60.4%; Pred. No. 0.11;
Matches 67; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 398 ATCAGGGATCCCGCGCTCTCACTGGGATAACGGGACCGCGGAGACTGACGCTTCAGC 457
Db 10181 ATCTGCAATACCACCGCTCATAGTACCAACCGGGGTCTCCACACAGCGCCTCAAG 10122
QY 458 CAGTACCATACCAAGCTTCATTTCCGAAGGACGACCAACCACTGCGC 508
Db 10121 CGAAGCGATCCCGAACCCCTCATAGTGAAGCTGCAGCACCACTGCGC 10071

RESULT 6

US-10-270-878-1/c
; Sequence 1, Application US/10270878
; Publication No. US20030082790A1
; GENERAL INFORMATION:

```

; APPLICANT: Sigrður Hjorleifsdóttir
; APPLICANT: Guðmundur O. Hreggvidsson
; APPLICANT: Ólafur H. Fridjónsson
; APPLICANT: Arnthor Aevársson
; APPLICANT: Jakob K. Kristjánsson
; TITLE OF INVENTION: Bacteriophage RM378 of a Thermophilic
; TITLE OF INVENTION: Host Organism
; FILE REFERENCE: 2739.1001-001
; CURRENT APPLICATION NUMBER: US/10/270,878
; PRIOR FILING DATE: 2002-10-11
; PRIOR APPLICATION NUMBER: US/09/585,858
; PRIOR FILING DATE: 2000-12-18
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 129908
; TYPE: DNA
; ORGANISM: Bacteriophage RM378
US-10-270-878-1

Query Match      3.4%; Score 40.6; DB 9; Length 129908;
Best Local Similarity 60.4%; Pred. No. 0.11;
Matches 67; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 398 ATCAGGATCCCGCGCTCTCACTGGCGATACGGGCGGCGGAGCTGACGCTTCAGC 457
DB 10181 ATCTGCAATACCAACCGCTCTAGTAGCAACACCGGGCTCCACACAGCGCGCTCAAG 10122

QY 458 CAGTACCATACCAACCGCTCTATTTCCGAAGGATGACACACACTGGC 508
DB 10121 CGAAGCGATCCCGAACCCCTATTAGATGAAGCTGCAGCACCATACATCGGC 10071

RESULT 7
US-10-270-786-1/c
; Sequence 1, Application US/10270786
; Publication No. US20030087392A1
; GENERAL INFORMATION:
; APPLICANT: Sigrður Hjorleifsdóttir
; APPLICANT: Guðmundur O. Hreggvidsson
; APPLICANT: Ólafur H. Fridjónsson
; APPLICANT: Arnthor Aevársson
; APPLICANT: Jakob K. Kristjánsson
; TITLE OF INVENTION: Bacteriophage RM378 of a Thermophilic
; TITLE OF INVENTION: Host Organism
; FILE REFERENCE: 2739.1001-001
; CURRENT APPLICATION NUMBER: US/10/270,786
; PRIOR FILING DATE: 2002-10-11
; PRIOR APPLICATION NUMBER: US/09/585,858
; PRIOR FILING DATE: 2000-12-18
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 129908
; TYPE: DNA
; ORGANISM: Bacteriophage RM378
US-10-270-786-1

Query Match      3.4%; Score 40.6; DB 9; Length 129908;
Best Local Similarity 60.4%; Pred. No. 0.11;
Matches 67; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 398 ATCAGGATCCCGCGCTCTCACTGGCGATACGGGCGGCGGAGCTGACGCTTCAGC 457
DB 10181 ATCTGCAATACCAACCGCTCTAGTAGCAACACCGGGCTCCACACAGCGCGCTCAAG 10122

QY 458 CAGTACCATACCAACCGCTCTATTTCCGAAGGATGACACACACTGGC 508
DB 10121 CGAAGCGATCCCGAACCCCTATTAGATGAAGCTGCAGCACCATACATCGGC 10071

RESULT 8
US-10-270-710-1/c
; Sequence 1, Application US/10270710
; Publication No. US20030092128A1
; GENERAL INFORMATION:
; APPLICANT: Sigrður Hjorleifsdóttir
; APPLICANT: Guðmundur O. Hreggvidsson
; APPLICANT: Ólafur H. Fridjónsson
; APPLICANT: Arnthor Aevársson
; APPLICANT: Jakob K. Kristjánsson
; TITLE OF INVENTION: Bacteriophage RM378 of a Thermophilic
; TITLE OF INVENTION: Host Organism
; FILE REFERENCE: 2739.1001-001
; CURRENT APPLICATION NUMBER: US/10/270,710
; PRIOR FILING DATE: 2002-10-11
; PRIOR APPLICATION NUMBER: US/09/585,858
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: 60/137,120
; PRIOR FILING DATE: 1999-06-02
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 129908
; TYPE: DNA
; ORGANISM: Bacteriophage RM378
US-10-270-710-1

Query Match      3.4%; Score 40.6; DB 9; Length 129908;
Best Local Similarity 60.4%; Pred. No. 0.11;
Matches 67; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 398 ATCAGGATCCCGCGCTCTCACTGGCGATACGGGCGGCGGAGCTGACGCTTCAGC 457
DB 10181 ATCTGCAATACCAACCGCTCTAGTAGCAACACCGGGCTCCACACAGCGCGCTCAAG 10122

QY 458 CAGTACCATACCAACCGCTCTATTTCCGAAGGATGACACACACTGGC 508
DB 10121 CGAAGCGATCCCGAACCCCTATTAGATGAAGCTGCAGCACCATACATCGGC 10071

RESULT 9
US-10-270-859-1/c
; Sequence 1, Application US/10270859
; Publication No. US20030092134A1
; GENERAL INFORMATION:
; APPLICANT: Sigrður Hjorleifsdóttir
; APPLICANT: Guðmundur O. Hreggvidsson
; APPLICANT: Ólafur H. Fridjónsson
; APPLICANT: Arnthor Aevársson
; APPLICANT: Jakob K. Kristjánsson
; TITLE OF INVENTION: Bacteriophage RM378 of a Thermophilic
; TITLE OF INVENTION: Host Organism
; FILE REFERENCE: 2739.1001-001
; CURRENT APPLICATION NUMBER: US/10/270,859
; PRIOR FILING DATE: 2002-10-11
; PRIOR APPLICATION NUMBER: US/09/585,858
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: 60/137,120
; PRIOR FILING DATE: 1999-06-02
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 129908
; TYPE: DNA
; ORGANISM: Bacteriophage RM378
US-10-270-859-1

Query Match      3.4%; Score 40.6; DB 9; Length 129908;
Best Local Similarity 60.4%; Pred. No. 0.11;
Matches 67; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 398 ATCAGGATCCCGCGCTCTCACTGGCGATACGGGCGGCGGAGCTGACGCTTCAGC 457
DB 10181 ATCTGCAATACCAACCGCTCTAGTAGCAACACCGGGCTCCACACAGCGCGCTCAAG 10122

QY 458 CAGTACCATACCAACCGCTCTATTTCCGAAGGATGACACACACTGGC 508
DB 10121 CGAAGCGATCCCGAACCCCTATTAGATGAAGCTGCAGCACCATACATCGGC 10071
```

```

; APPLICANT: Sigrður Hjorleifsdóttir
; APPLICANT: Guðmundur O. Hreggvidsson
; APPLICANT: Ólafur H. Fridjónsson
; APPLICANT: Arnthor Aevársson
; APPLICANT: Jakob K. Kristjánsson
; TITLE OF INVENTION: Bacteriophage RM378 of a Thermophilic
; TITLE OF INVENTION: Host Organism
; FILE REFERENCE: 2739.1001-001
; CURRENT APPLICATION NUMBER: US/10/270,878
; PRIOR FILING DATE: 2002-10-11
; PRIOR APPLICATION NUMBER: US/09/585,858
; PRIOR FILING DATE: 2000-12-18
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 129908
; TYPE: DNA
; ORGANISM: Bacteriophage RM378
US-10-270-878-1

Query Match      3.4%; Score 40.6; DB 9; Length 129908;
Best Local Similarity 60.4%; Pred. No. 0.11;
Matches 67; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 398 ATCAGGATCCCGCGCTCTCACTGGCGATACGGGCGGCGGAGCTGACGCTTCAGC 457
DB 10181 ATCTGCAATACCAACCGCTCTAGTAGCAACACCGGGCTCCACACAGCGCGCTCAAG 10122

QY 458 CAGTACCATACCAACCGCTCTATTTCCGAAGGATGACACACACTGGC 508
DB 10121 CGAAGCGATCCCGAACCCCTATTAGATGAAGCTGCAGCACCATACATCGGC 10071

RESULT 7
US-10-270-786-1/c
; Sequence 1, Application US/10270786
; Publication No. US20030087392A1
; GENERAL INFORMATION:
; APPLICANT: Sigrður Hjorleifsdóttir
; APPLICANT: Guðmundur O. Hreggvidsson
; APPLICANT: Ólafur H. Fridjónsson
; APPLICANT: Arnthor Aevársson
; APPLICANT: Jakob K. Kristjánsson
; TITLE OF INVENTION: Bacteriophage RM378 of a Thermophilic
; TITLE OF INVENTION: Host Organism
; FILE REFERENCE: 2739.1001-001
; CURRENT APPLICATION NUMBER: US/10/270,786
; PRIOR FILING DATE: 2002-10-11
; PRIOR APPLICATION NUMBER: US/09/585,858
; PRIOR FILING DATE: 2000-12-18
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 129908
; TYPE: DNA
; ORGANISM: Bacteriophage RM378
US-10-270-786-1

Query Match      3.4%; Score 40.6; DB 9; Length 129908;
Best Local Similarity 60.4%; Pred. No. 0.11;
Matches 67; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 398 ATCAGGATCCCGCGCTCTCACTGGCGATACGGGCGGCGGAGCTGACGCTTCAGC 457
DB 10181 ATCTGCAATACCAACCGCTCTAGTAGCAACACCGGGCTCCACACAGCGCGCTCAAG 10122

QY 458 CAGTACCATACCAACCGCTCTATTTCCGAAGGATGACACACACTGGC 508
DB 10121 CGAAGCGATCCCGAACCCCTATTAGATGAAGCTGCAGCACCATACATCGGC 10071
```

Db 10181 ATCTGCAATACCAACCGTCTAGTAGCAACCAACCGGGCTCCACACAGGACGCTCAAG 10122
 QY 458 CAGTACATACCAACCGTCTATTTCCGAGGATGACACACACTGGC 508
 Db 10121 CGAAGCATCCGAACCGCTCATTAGTAGAAGCTGCAGCACCATCGGC 10071

RESULT 10

US-09-974-300-1730/c
 ; Sequence 1730, Application US/09974300
 ; Patent No. US20020146721A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Berka, Randy M.
 ; APPLICANT: Clausen, Ib Groth
 ; TITLE OF INVENTION: Methods For Monitoring Multiple Gene
 ; FILE OF INVENTION: Expression
 ; FILE REFERENCE: 10085.500-US
 ; CURRENT APPLICATION NUMBER: US/09/974,300
 ; CURRENT FILING DATE: 2001-10-05
 ; PRIOR APPLICATION NUMBER: 09/680,598
 ; PRIOR FILING DATE: 2000-10-06
 ; PRIOR APPLICATION NUMBER: 60/279,526
 ; PRIOR FILING DATE: 2001-03-27
 ; NUMBER OF SEQ ID NOS: 8481
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 1730
 ; LENGTH: 1108
 ; TYPE: DNA
 ; ORGANISM: Bacillus licheniformis
 US-09-974-300-1730

Query Match 3.1%; Score 37; DB 10; Length 1108;
 Best Local Similarity 54.9%; Pred. No. 0.09;
 Matches 73; Conservative 0; Mismatches 60; Indels 0; Gaps 0;
 QY 387 TTCTGAATGATACATGAGGATCCCGCGCTCTCACTGGCGATACGCGGCGCGGAGACT 446
 Db 926 TTCTGTATAAATCTTTCAGCGCGCGGAGTCTGCCCGATACAGGCGTCCCGCAGCG 867
 QY 447 GAGCTTCACGCCAGTACCAACAGCTTCTTCCGAGGATGACCAACACTG 506
 Db 866 AGCGCTTCAAGGCGGAAATTTCCGAGGTTTCGTCGGAGAAGGAAGCAACAATCT 807
 QY 507 GCAATCCGCTAGA 519
 Db 806 GAGTCGCATAGA 794

RESULT 11

US-09-825-561A-7
 ; Sequence 7, Application US/09825561A
 ; Patent No. US20020137677A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Sprecher, Cindy A.
 ; APPLICANT: West, James W.
 ; APPLICANT: Presnell, Scott R.
 ; APPLICANT: Holly, Richard D.
 ; APPLICANT: Nelson, Andrew J.
 ; TITLE OF INVENTION: SOLUBLE ZALPHAL1 CYTOKINE RECEPTORS
 ; FILE REFERENCE: 00-22
 ; CURRENT APPLICATION NUMBER: US/09/825,561A
 ; CURRENT FILING DATE: 2000-04-05
 ; PRIOR APPLICATION NUMBER: US 60/194,731
 ; PRIOR FILING DATE: 2000-04-05
 ; PRIOR APPLICATION NUMBER: US 60/222,121
 ; PRIOR FILING DATE: 2000-07-28
 ; NUMBER OF SEQ ID NOS: 86
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 7
 ; LENGTH: 654
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence

; FEATURE:
 ; OTHER INFORMATION: Degenerate polynucleotide sequence of soluble
 ; OTHER INFORMATION: zalphal1 Receptor polypeptide as shown in SEQ ID
 ; NAME/KEY: misc_feature
 ; LOCATION: (1)_(654)
 ; OTHER INFORMATION: n = A,T,C or G
 US-09-825-561A-7
 Query Match 3.1%; Score 36.6; DB 10; Length 654;
 Best Local Similarity 30.7%; Pred. No. 0.089;
 Matches 74; Conservative 35; Mismatches 132; Indels 0; Gaps 0;
 QY 883 GTGAGATAGGCACAAATCAATCAACAGATCAATTCAGGAAAACGTTCTGCTTTAC 942
 Db 60 GTGGAAYTTCAYCCNWSNACNTNACNTNACNTGCGARGAYCARTAYGARGARYTNA 119
 QY 943 GGCTGATAGTGGTTTCTGCTGACATAGTGAAGCGGTGCACAGCATATCAGACGGCTCA 1002
 Db 120 RGATGARGCNACNWSNTGYWSNTNCAIYNGWSNCAAYAGCNACNACNTA 179
 QY 1003 GTCTGTATATTAATCTGTCATGGCCACTATGGCAGATGACAGATCAGGTTTAAATCCC 1062
 Db 180 YACNTGYCATGAGCTNTTCAYTTATGGCNGAYGAYATHTTYSNGTNAAYATAC 239
 QY 1063 CGATAATCCCTCGAAGTCTGAGGATGAAGAGGTTGAAGGCTTCTCTCTGAAAGGAATA 1122
 Db 240 NGAYCARWSNGNAAATAYNSNCARGARTGYGWNSTNTTYTNTGNCARWSNATHAA 299
 QY 1123 A 1123
 Db 300 R 300
 RESULT 12
 US-09-864-761-14748/c
 ; Sequence 14748, Application US/09864761
 ; Patent No. US20020048763A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Penn, Sharon G.
 ; APPLICANT: Hanzel, David K.
 ; APPLICANT: Chen, Wensheng
 ; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
 ; FILE REFERENCE: Aeomica-X-1
 ; CURRENT APPLICATION NUMBER: US/09/864,761
 ; CURRENT FILING DATE: 2001-05-23
 ; PRIOR APPLICATION NUMBER: US 60/180,312
 ; PRIOR FILING DATE: 2000-02-04
 ; PRIOR APPLICATION NUMBER: US 60/207,456
 ; PRIOR FILING DATE: 2000-05-26
 ; PRIOR APPLICATION NUMBER: US 09/632,366
 ; PRIOR FILING DATE: 2000-08-03
 ; PRIOR APPLICATION NUMBER: GB 24263.6
 ; PRIOR FILING DATE: 2000-10-04
 ; PRIOR APPLICATION NUMBER: US 60/236,359
 ; PRIOR FILING DATE: 2000-09-27
 ; PRIOR APPLICATION NUMBER: PCT/US01/00666
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00667
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00664
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00669
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00665
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00668
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00663
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00662

```

; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 14748
; LENGTH: 598
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC018525.4
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.73
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.62
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.48
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.66
; US-09-864-761-14748

Query Match          3.1%; Score 36.2; DB 10; Length 598;
Best Local Similarity 56.2%; Pred. No. 0.12;
Matches 68; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

QY 1053 TTAATTCCTCCGATATCGTGAAGCTGAGAGTGAAGAGGAGTGAAGGCTGTTCCTG 1112
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1113 AAAGGAATAAAGTGCATCATGCGCTCTTTTCTGGCTCCGGAGCAATTTACTTTT 1172
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 443 ACAGAAATAAAGTGCATCATGCGCTCTTTTCTGGCTCCGGAGCAATTTACTTTT 1172
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1173 T 1173
Db      383 T 383

RESULT 13
US-09-825-561A-66
; Sequence 66, Application US/09825561A
; Patent No. US20020137677A1
; GENERAL INFORMATION:
; APPLICANT: Sprecher, Cindy A.
; APPLICANT: West, James W.
; APPLICANT: Presnell, Scott R.
; APPLICANT: Holly, Richard D.
; APPLICANT: Nelson, Andrew J.
; TITLE OF INVENTION: SOLUBLE ZALPHA11 CYTOKINE RECEPTORS
; FILE REFERENCE: 00-22
; CURRENT APPLICATION NUMBER: US/09/825,561A
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/194,731
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/222,121
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 86
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 66
; LENGTH: 741
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Degenerate polynucleotide sequence of SEQ ID NO:69
; NAME/KEY: misc_feature
; LOCATION: (1)...(741)
; OTHER INFORMATION: n = A,T,C or G
US-09-825-561A-66
```

```

Query Match          3.1%; Score 36.2; DB 10; Length 741;
Best Local Similarity 27.6%; Pred. No. 0.13;
Matches 83; Conservative 48; Mismatches 170; Indels 0; Gaps 0;

QY 833 AACCGTTACGGTGTCTCTGACACCTTCCGCCATFACATGCGGCATCATCGGTGAGATAGG 892
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 139 AAYTTNCAAYCCNNSNACNTYNACNTGNACNTGATHYTNWSNAAAYAYACNGGNTGYTAY 198
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 893 CACAACAATGAATACACAGATAAATTCAGGGAACAGTCTCTGCTTACGGGTGATGTA 952
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 199 ATHAARGAYMGNACNTYNGAYTTNMGNCARGAYCARTAYGARGARTYNAARGAYGARGCN 258
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 953 GGTTTTGTCTGACAAATAGTGAAGCGGTGACAGCATATCAGACGGCTCAGTCTCTCTAT 1012
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 259 ACNWSNTGYWSNTNCAYMNGWSNCAAYAGCNACNCAAYACNACNTAYACNTGYCAY 318
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1013 ATTACTGTCTATGCCACTATGGCAGATGACCATCATCAGGTTTAAATCCCGGATATCCG 1072
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 319 ATGGAYGTNTTYCAVTTTATGCGNGAYGAYATHTTYSNNGTNAAYATHACNGAYCARWSN 378
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1073 TCGAAGTCTGAGGATGAAGGAAGGTGAAGGCTGTTCCTGAAAGGAATAAAAGTGACATC 1132
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 379 GGNAAATAYWSNCARGARTGGNWSNTTYYTNTNCCNARGWSNMGNCARTAYAAVATH 438
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1133 A 1133
Db      439 W 439

RESULT 14
US-09-974-300-757/c
; Sequence 757, Application US/09974300
; Patent No. US20020146721A1
; GENERAL INFORMATION:
; APPLICANT: Berkla, Randy M.
; APPLICANT: Clausen, Ib Groth
; TITLE OF INVENTION: Methods For Monitoring Multiple Gene
; FILE REFERENCE: 10085.500-US
; CURRENT APPLICATION NUMBER: US/09/974,300
; CURRENT FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 09/680,598
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/279,526
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 8481
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 757
; LENGTH: 1427
; TYPE: DNA
; ORGANISM: Bacillus licheniformis
US-09-974-300-757

Query Match          3.0%; Score 36; DB 10; Length 1427;
Best Local Similarity 46.0%; Pred. No. 0.24;
Matches 160; Conservative 0; Mismatches 185; Indels 3; Gaps 1;

QY 398 ATCAGGATCCGCCGCCGTCTCCTGCGCATACGGGACGCGGAGACTGACGCTCAGC 457
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1215 ATCATGAAGCCGCCCTGTTTCCCTGACAAACCGGACCGCGCGTACCTGAGGGCAATCAG 1156
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 458 CAGTACCATACCAACAGCTTCATTTTCGAAGGCATACCAACACACCTGCGCAATCCGTA 517
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1155 CTGTCCAGACCGCACGCTCGAATTTTGAAGGCATCAGGAAGATCGCTCGCGGCATA 1096
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 518 GACCGGTAACGCTGGGAAAGGACCTGCCATTAACACATCTCCGCTCATTTCCAGGTG 577
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1095 AATCTTGGCGTGAAGGCTTCATGAAACCGATATACGCTTGCACCTGTCCGGAAGCA 1036
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 578 TTCTGTCTGCTACGCGACGAGCTGCTTCGTATTTTACGCGCGCGCCACACACGAGCA 637
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1035 ATGCTCCATATAGCGGAGTAATTTTCAATTTCCGCTTCTCCGGTCTCTAAACAGTACG 976
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

QY 638 CGGAATGATTCCTCCCTTCACCTGATACATAACAGCAGCAGCAATAATTCATGTC 697
Db 975 CTGATGCTCTCTCTTCA--ACAGTTATGCAATACCGGCTGTGATCAATCGAAGCC 919
QY 698 TTTTCGGGAGTACATCCACCTGAAGCAAGCGGAACAATTGTC 745
Db 918 TTTTGTCTGTCAGCCTTCCACCATGCTGATGAGCGGAATGTCCTTC 871

RESULT 15
US-09-918-995-8152/c
; Sequence 1852, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8152
; LENGTH: 480
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(480)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-8152

Query Match 3.0%; Score 35.4; DB 9; Length 480;
Best Local Similarity 54.1%; Pred. No. 0.19;
Matches 72; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

QY 1013 ATTACTGTCATGGCCATATGGCAGATGACCATGAGTTTAAATCCCGATATCCG 1072
Db 429 ATCTCTGACATGGCCATATGAAGATCTCTCGAAGATCCCACTCCAGAGATTG 370

QY 1073 TCGAAGTCTGAGGATGAAGAGTGAAGCTGCTCTCTGAAGCAATAAAGTACATC 1132
Db 369 CTTAAACAGAGTTCATGAAGCTTGATGATGATGATGATGATGATGATGATGAT 310

QY 1133 ATGCCCTCTTTT 1145
Db 309 ATTTCCTTTGTT 297

RESULT 16
US-09-974-300-1702/c
; Sequence 1702, Application US/09974300
; Patent No. US20020146721A1
; GENERAL INFORMATION:
; APPLICANT: Berka, Randy M.
; APPLICANT: Clausen, Ib Groth
; TITLE OF INVENTION: Methods For Monitoring Multiple Gene
; FILE REFERENCE: Expression
; CURRENT APPLICATION NUMBER: US/09/974,300
; CURRENT FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 09/680,598
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/279,526
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 8481
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1702
; LENGTH: 1135
; TYPE: DNA

; ORGANISM: Bacillus licheniformis
US-09-974-300-1702

Query Match 3.0%; Score 35; DB 10; Length 1135;
Best Local Similarity 50.9%; Pred. No. 0.46;
Matches 83; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

QY 352 CATCCCCCAGGGGTAAACAGCGTCCCTGTCTACATCTTCTGAATGATCAGGGATCCCGC 411
Db 967 CGTCTCCCTCGGGAACCAAAAATCCGTTACGCGCTTGAATGACTTCCCGGTGTCGC 908

QY 412 CCGTCTCACTGGCGATAACGGCGACGCGGAGAGCTGACGCTTACGCCAGTACCATACCA 471
Db 907 GGTGTCCTGCTGCAACGCGGCGCTTCTGAGCCATCTCTCAAGGAGGTTCATGCCGA 848

QY 472 AGCTTTCATTTCCGAAGCATACACCATCTGCGCAATCCG 514
Db 847 GCCTTCCCTGATGCTTGGAGCGACACACGCTGCAAGCTG 805

RESULT 17
US-10-123-155-536/c
; Sequence 536, Application US/10123155
; Publication No. US20030068794A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tamas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C30
; CURRENT APPLICATION NUMBER: US/10/123,155
; CURRENT FILING DATE: 2002-04-15
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 536
; LENGTH: 513
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-123-155-536

Query Match 2.9%; Score 34.8; DB 9; Length 513;
Best Local Similarity 6.9%; Pred. No. 0.32;
Matches 30; Conservative 115; Mismatches 230; Indels 0; Gaps 0;

QY 303 CGTTCGGTGGGATAAAAATCCGCGAGTCCGCGCTCCATCGACACATCCCCCAGC 362
Db 459 YK.SY.....MKST.RYWSBR.B...B.BS...TBSCT..MD.S.D.THHA..C..ATC 400

QY 363 GGTAACAGCGTCCCTGTCTGATGATGATGATGATGATGATGATGATGATGATGATG 422
Db 399 BA.....MS.DM.Y..M.A.M..TYT.H.Y.AASWCG.SRNBETH.BNEC....M.STG 340

QY 423 GCGATAACGGCGCGCGGAGACTGACGCTTACGCCAGTACCATACCAACGCTTCATTT 482
Db 339 ..AN.Y.CM.SM.WNB.SG.NYSG.W.NCT.S.HN.S.WSH....C..A.MNSH.N.Y.. 280

QY 483 TCCGAGGCGATGACCCACCATCGGCTAGACCGGTACCGGTGCGGAAGGCA 542

Db 279 N.BG..BS:CS.T...NCS.H.Y...SSWASKA.C.BS.MNW...R.N...SB.Y...T.N. 220
 QY 543 CTOCCATTACACATCCCGCTCAATVCCAGGTGTTCTGCTGACGACGACGCTGTT 602
 Db 219 MS..ND..D.M.Y.KCT.BNVT.SY.YNRC.H....NYGH...YB..AY.SNSY.D.S. 160
 QY 603 TCGTATCTTACGCGCGCGCCACACGAGCCAGCGGAATATCTCCCTTCATCTTC 662
 Db 159 .MY.CY...SC.SD..NRS.H.NT.NAB.Y.ANN..RS.YNSS....M.YY.CN.TN.H. 100
 QY 663 AGCTGATACATACACGACGACATAAATCTGCTCTTTTCGGGACGTAGCATCCCAACC 722
 Db 99 NS.DNDH.R.WA..N.CM..NRM.M..SNRY.S.C.GCTS.SS.....M.S.GRBKMC.G 40
 QY 723 TGAAGGATAACGGGA 737
 Db 39 YGCM.GCY.TSS.KA 25

RESULT 18

US-10-184-644-312/c
 ; Sequence 312, Application US/10184644
 ; Publication No. US20030044930A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Baker, Kevin P.
 ; APPLICANT: Chen, Jian
 ; APPLICANT: Desnoyers, Luc
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Pan, James
 ; APPLICANT: Smith, Victoria
 ; APPLICANT: Watanabe, Colin K.
 ; APPLICANT: Wood, William I.
 ; APPLICANT: Zhang, Zemin
 ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 ; FILE REFERENCE: P3430R1C227
 ; CURRENT APPLICATION NUMBER: US/10/184,644
 ; CURRENT FILING DATE: 2002-06-28
 ; Prior Application removed - See File Wrapper or Palm
 ; NUMBER OF SEQ ID NOS: 612
 ; SEQ ID NO 312
 ; LENGTH: 802
 ; TYPE: PRT
 ; ORGANISM: Homo Sapien
 US-10-184-644-312

Query Match 2.9%; Score 34.8; DB 9; Length 802;
 Best Local Similarity 5.2%; Pred. No. 0.43;
 Matches 28; Conservative 170; Mismatches 336; Indels 0; Gaps 0;
 QY 190 TGTGTCATTTCCCTTAGCCTGACCTAGCCAGACACAAATGATCTGCGGTTCTGTAA 249
 Db 591 YRHYK..B.MTMH.NSHSN...BB.MBN.N...GAM..NCK.YB.KH.YCS.....N.MM 532
 QY 250 TATCAAACCGGTACTCAATATCTCTGCGCTGGCTGCCATCATCCGGAAGCGTTCCG 309
 Db 531 HM..MD.SMSC.....B.MM.SNS..S.SCS..N.YMT...TH...D.B...MS.A...R 472
 QY 310 GTCGGGATAAATAATCCGCGAGTCCGCGGTCTGACATGACACATCCCGCGGTAACA 369
 Db 471 TDYBY.BAKMCMMDMR....H..R.TCRK.NBYRCR...M.H..YCYAMTNTN.K.KB 412
 QY 370 GCGTCCCTGTACATTTCTTGAATGACATCAGGATCCGCGGTCTCACTGGCGATAA 429
 Db 411 .KB.HH..T.Y.TSHKYY...CHBTSR.RMCA...W.....Y.Y...YK.S...C...H. 352
 QY 430 CGGCGACGCGGAGACTGACGCTTACCGAGTACCATACCAACGCTTCTTCCGAAG 489
 Db 351 MD..ANSRMSK.B..YD...AYDY..S.MYS.KRSD.YR.R..NMN.D.A.T.D.MSND. 292
 QY 490 GCATGACACACACTGGCAATCCGTTAGCCGTAACCGGTGGAAGGACCGACCTGCCA 549
 Db 291 HY.RCMWN..RN.R...YK.RSWBG.BCT.YYBYCB.BH.DAARK.Y.GMC.D.BKYY. 232

Db 291 HY.RCMWN..RN.R...YK.RSWBG.BCT.YYBYCB.BH.DAARK.Y.GMC.D.BKYY. 232
 QY 550 TTAACACATCTCCGCTCATTCCTCCAGGTGTTCTGCTGACGACGAGCGTCTGCTATT 609
 Db 231 B.YSK.BC.CCKG.N.C...T.MCK..AAC.CA.C....SSN.SY..N..YHCM.RBHH 172
 QY 610 CTTACGCGCGCGCGCCACACGAGCCAGCAAAATGATTCCTTCCATCTTCAGCTGAT 669
 Db 171 HHYK.W.RMH.RDHDKM.KK.SMM...RSHEBCY..BB...BSAHSK.SS....BMC..A 112
 QY 670 ACAATACAGCGAGCAATAATCATCTGCTCTTTTCGGGACGTAGCATCCCACT 723
 Db 111 MSWAYRTTBYA..RM.TAKCB...NHY.CCH.HSC..W..TCYTHCYTCT 58

RESULT 19

US-10-184-634-312/c
 ; Sequence 312, Application US/10184634
 ; Publication No. US20030068694A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Baker, Kevin P.
 ; APPLICANT: Chen, Jian
 ; APPLICANT: Desnoyers, Luc
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Pan, James
 ; APPLICANT: Smith, Victoria
 ; APPLICANT: Watanabe, Colin K.
 ; APPLICANT: Wood, William I.
 ; APPLICANT: Zhang, Zemin
 ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 ; FILE REFERENCE: P3430R1C217
 ; CURRENT APPLICATION NUMBER: US/10/184,634
 ; CURRENT FILING DATE: 2002-06-28
 ; Prior Application removed - See File Wrapper or Palm
 ; NUMBER OF SEQ ID NOS: 612
 ; SEQ ID NO 312
 ; LENGTH: 802
 ; TYPE: PRT
 ; ORGANISM: Homo Sapien
 US-10-184-634-312

Query Match 2.9%; Score 34.8; DB 9; Length 802;
 Best Local Similarity 5.2%; Pred. No. 0.43;
 Matches 28; Conservative 170; Mismatches 336; Indels 0; Gaps 0;
 QY 190 TGTGTCATTTCCCTTAGCCTGACCTAGCCAGACACAAATGATCTGCGGTTCTGTAA 249
 Db 591 YRHYK..B.MTMH.NSHSN...BB.MBN.N...GAM..NCK.YB.KH.YCS.....N.MM 532
 QY 250 TATCAAACCGGTACTCAATATCTCTGCGCTGGCTGCCATCATCCGGAAGCGTTCCG 309
 Db 531 HM..MD.SMSC.....B.MM.SNS..S.SCS..N.YMT...TH...D.B...MS.A...R 472
 QY 310 GTCGGGATAAATAATCCGCGAGTCCGCGGTCTGACATGACACATCCCGCGGTAACA 369
 Db 471 TDYBY.BAKMCMMDMR....H..R.TCRK.NBYRCR...M.H..YCYAMTNTN.K.KB 412
 QY 370 GCGTCCCTGTACATTTCTTGAATGACATCAGGATCCGCGGTCTCACTGGCGATAA 429
 Db 411 .KB.HH..T.Y.TSHKYY...CHBTSR.RMCA...W.....Y.Y...YK.S...C...H. 352
 QY 430 CGGCGACGCGGAGACTGACGCTTACCGAGTACCATACCAACGCTTCTTCCGAAG 489
 Db 351 MD..ANSRMSK.B..YD...AYDY..S.MYS.KRSD.YR.R..NMN.D.A.T.D.MSND. 292
 QY 490 GCATGACACACACTGGCAATCCGTTAGCCGTAACCGGTGGAAGGACCGACCTGCCA 549
 Db 291 HY.RCMWN..RN.R...YK.RSWBG.BCT.YYBYCB.BH.DAARK.Y.GMC.D.BKYY. 232


```

; APPLICANT: Wood,William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330R1C30
; CURRENT APPLICATION NUMBER: US/10/123,155
; CURRENT FILING DATE: 2002-04-15
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 96
; LENGTH: 716
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-123-155-96

Query Match          2.8%   Score 33.2; DB 9; Length 716;
Best Local Similarity 8.0%; Pred.No.1.4;
Matches 40; Conservative 135; Mismatches 322; Indels 0; Gaps 0;

Qy      343 ATGCAGACATCCCCCAGGGTAACAGCGTCCCTGTGCATATTTCTTCTGAATGACATCAG 402
Db      :| : : : : :| : : : : :| : : : : :| : : : : :| : : : : :| : : : : :|
Qy      541 RTA.AMNMNH..BSYV..CK..MHMS....MKABK.MCHBTM..TT.NB.N...TCS.CB 482
Db      :| : : : : :| : : : : :| : : : : :| : : : : :| : : : : :| : : : : :|
Qy      403 GGATCCGCGCGTCTACTGGCGATAACGGCAGCGCGAGACTGACGCTTCACGCCAGTA 462
Db      | : : : : :| : : : : :| : : : : :| : : : : :| : : : : :| : : : : :|
Qy      481 ACYT....HBS.H.M.MN.RWM...AYCMBH.Y.BAB.KH.YC.Y..T.NH.T...NATC. 422
Db      | : : : : :| : : : : :| : : : : :| : : : : :| : : : : :| : : : : :|
Qy      463 CCATACCAACGGCTTCATTTTCCGAAGCGATGACCACCACCACTGGCAATCGGGTAGACGG 522
Db      | : : : : :| : : : : :| : : : : :| : : : : :| : : : : :| : : : : :|
Qy      421 ..BC.N.M.CHK.T...N.A.YSR.DK.S..YMCBSH..H...BG.TNTMT.Y..NY.Y 362
Db      | : : : : :| : : : : :| : : : : :| : : : : :| : : : : :| : : : : :|
Qy      523 GTAACGCTGGGAAAGGCGACCTCCCATTAACACATCTCGCTCATTCGCCAGGGTTCGTG 582
Db      | : : : : :| : : : : :| : : : : :| : : : : :| : : : : :| : : : : :|
Qy      361 STCECB.K.H..SCSTRR..BHT..CTBTYT..AMCAC..CB...CM..MCC..A.M..NM 302
Db      | : : : : :| : : : : :| : : : : :| : : : : :| : : : : :| : : : : :|
Qy      583 TCTGCTGACGCAGCAGCGTCTCGTATCTTCACCGCGGGCGCCACCACGAGCCAGCGAA 642
Db      301 ...BB.....MT..BCEMD..ABNWK.B.HBTSH.CAAAY.YBS...NM..CR.C....B 242
Qy      643 ATGATTTCCCTTCATCTTCAGCTGATACATAACAGCAGCATAAATTCATGCTCTTTT 702
Db      :| : : : : :| : : : : :| : : : : :| : : : : :| : : : : :| : : : : :|
Qy      241 ...YAYY.SHNAM.A.T.M..C.T.CAM.TH.DT...NMAMK.MH.S.BHSCYHY... 182
Db      :| : : : : :| : : : : :| : : : : :| : : : : :| : : : : :| : : : : :|
Qy      703 CGGAGCTAGCATCCCCACCTGACGATAAGCGGAACATTCCTGCTGATGCACCCGAGG 762
Db      | : : : : :| : : : : :| : : : : :| : : : : :| : : : : :| : : : : :|
Qy      181 CMB..S.T.N.A.T.Y.SAS...YYAYSM.AM.CYS...B.WR..HSG.S..TYSDDYR.GS 122
Db      :| : : : : :| : : : : :| : : : : :| : : : : :| : : : : :| : : : : :|
Qy      763 CGTGGATGATGAGGGGTAAAGCGTGGCATGCGCTTCATTATGCAATGCGGGCGAGTCAAC 822
Db      :| : : : : :| : : : : :| : : : : :| : : : : :| : : : : :| : : : : :|
Qy      121 SY.AS..H.NCRMN....TSBDSADWDSS.NWCMG.C...N.B..C.M..S.C.HY 62
Db      :| : : : : :| : : : : :| : : : : :| : : : : :| : : : : :| : : : : :|
Qy      823 CCGGTGGAATAACGGTT 839
Db      :| : : : : :| : : : : :| : : : : :| : : : : :| : : : : :| : : : : :|
Qy      61 .N.S..SS.T.D..BBH 45
Db      :| : : : : :| : : : : :| : : : : :| : : : : :| : : : : :| : : : : :|

```

RESULT 26
US-10-184-644-348/c
; Sequence 348, Application US/10184644
; Publication No. US20030044930A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

```

; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: F3430R1C227
; CURRENT APPLICATION NUMBER: US/10/184,644
; CURRENT FILING DATE: 2002-06-28
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 348
; LENGTH: 777
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-184-644-348

Query Match      2.88; Score 33.2; DB 9; Length 777;
Best local Similarity 7.5%; Pred. No. 1.5;
Matches 42; Conservative 179; Mismatches 342; Indels

QY 177 ATCCGCGGTATCTGTGCTATTCCTCCCTTAGCTGACTAGCCAGACACACAA
      : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
DB 765 DDYNYMMWK..KDMWM.CCMNY.YYM.VDWK..GR.H.S.N.SS....RH
      : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
QY 237 GCGCTTGTGTAATATCAAACCGGTACTACTAATATCTTCTGGCGTGGCTG
      : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
DB 705 .HMB.C.D.TY.AN.K..BN.A.MB.AD..AD..TMRREKCSHM..
      : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
QY 297 CGGAAGCGTCCCGTGGGATATAAAATCCGCAGTGCCTGGGTCCATCGCA
      : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
DB 645 M..Y.H.M..YD.HCSY..RWM.AT..SM..G...ASN...C..BM.HTA
      : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
QY 357 CCCACGGGTAAACAGCGTCCCTGTGCATATCTTCTGAATGACATCAGGATG
      : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
DB 585 WG.A..HCRMBH.YYYYMSA.TRYSGTNCHWTGR.HYT.GGTGTWCRMH
      : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
QY 417 TCACTGCGGATAACGGGACGCGGAGACTGACGCTTCAGCCAGTACCATA
      : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
DB 525 CHYSC.R....M.S..KN...SSDM.....BB..KNWM.MS.SBBM.BAL
      : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
QY 477 TCATTTCCGAAGCATGACCACACCACTGGCAATCCGGTAGACCGGTAAC
      : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
DB 465 HR.CH.T.BDHHB..A.YRHHB.YM.A.CCTB.RBSMRKBSDYM..S.BHH
      : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
QY 537 AGGCACTCCCATTAACACATCTCCGCTCATTCGCCAGGTGTCTGCTGCTG
      : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
DB 405 .HRAMS.GAC.Y.R..YCHR.BWVDHTS.MDTR.CN.BTY.HPKSRBGTSS
      : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
QY 597 CGTGCTTCGTATCTTCACCCCGCGGCCACCAACGAGCCAGCGCAATGAT
      : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
DB 345 .BCRBB.NY.HYA...R.H...H.RAHTCHSC..SG..YTM..AAWMN..S
      : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
QY 657 ATCTTCAGCTGTACATACACACGACGACATAATTCATGCTCTTTTCGGG
      : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
DB 285 BYCBYS..AMHSASC...SS.Y...R.MHHH.NRAH.....AC..MTCN.WFD
      : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
QY 717 CCCACCTGAACGATAAGCGGAAC 739
      : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
DB 225 DHD.A.C.SYA.TAHMC..HSTAC 203
      : : : : : | : : : : : | : : : : : | : : : : : | : : : : :

```

RESULT 27
US-10-184-634-348/c
; Sequence 348, Application US/10184634
; Publication No. US2003008684A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin

```

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1C217
; CURRENT APPLICATION NUMBER: US/10/184,634
; CURRENT FILING DATE: 2002-06-28
; Prior Application removed - See file Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 346
; LENGTH: 777
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-184-634-348

```

Query Match	2.8%;	Score 33.2;	DB 9;	Length 777;
Best Local Similarity	7.5%;	pred. 0.1.5;		
Matches 42;	Conservative 179;	Mismatches 342;	Indels 0;	Gaps 0;
QY	177	ATCCGCCGTTATCTGTTGCAATTTCCCTTACCGCTGACTACGCCAGAGACACAATGATCTGT	236	
Db	765	DDYNYMMK..KDWMM.CCMNY.YYM.YDWK..GR.H.S.N.SS....RHMY.YS.T.706		
QY	237	GCGGTCCTGTAATATCAACCCGCTACTCAATATCTCTCTGCGCTGCGCTGCCATCATC	296	
Db	705	.HMB.C.D.IY.AN.K..N..BN.A.MB.AD..TMGRKCSHM..SY...CR.A.646		
QY	297	CGGAAGCGTTCGGTTCGGGATAAAAAATCGCGCAGTGCGCCGCTCCATCGACACATCC	356	
Db	645	M.Y.H.M...YD.HCSY..RWM.AT..SM..G...ASN...C..BM.HTA.DS.SH..H.586		
QY	357	CCACGGGTAAACAGCGTCCCTGTGCACATTTCTTGAATGACATCAGGAGTACCGCGCGTC	416	
Db	585	WG.A..HCRMBH.YTYMYSA.TRYSGTNCHWTGR.HYT.GGHTGTMCRAGVD.S..B.526		
QY	417	TCATCGGGATAACGGGCACGCCGAGACTGACGCTTCAGCCAGTACCATATCAACAACGCT	476	
Db	525	CHYC.R...M.S..KN....SSDM.....BB..KNWM.MS.SBBM.BAC.HAC..KB.466		
QY	477	TCATTTTCGGAAGGATGACACACACACTCGCAATCGGTAGACCGGTAAACGCTGGGAAA	536	
Db	465	HR.CH.T.BDHHB..A.YRHBN.YM.A.CCTB.RBSMRKBSDYM..S.BHH..HVASM...406		
QY	537	AGGGCACTGCCATTAAACATATCTCCGCTCATTTCCAGGTGTTCTCTGCTGACGCGAGA	596	
Db	405	.HRAMS.GAC.Y.R..YCHR.BWDHYS.MDTR.CN.BTY.HFKSRBGTSCM..SSAAA.346		
QY	597	CGTGCTTCGTATTTCTTCAGCCCGCGGCCACACGAGCAGCGAAGTATTCCTTCCTCC	656	
Db	345	.BCRBB.NY.HVA...R.H...H.RAHTCHSC..SG..YTM..AAWMN..SY.CCBHNG.286		
QY	657	ATCTCAGCTGATACATAACAGGAGATCAATTCATGCTCTTTTCGGGACGTAGCATC	716	
Db	285	BYCBYS..AMHSASC..SS.Y...R.MHHH.NRAH.....AC..MTCN.WRD.S.HAY.RD.226		
QY	717	CCACCTGACGATAAGCGGAAC	739	
Db	225	DHDA.C.SYA.TAHMC...HSTAC.203		

RESULT 28
US-10-123-155-306/c
; Sequence 306, Application US/10123155
; Publication No. US20030068794A1
; GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.
APPLICANT: Beresford, Maureen
APPLICANT: Beforge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao Wei-Qiang
APPLICANT: Gerritsen, Mary E
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.

```

; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330R1C30
; CURRENT APPLICATION NUMBER: US/10/123,155
; CURRENT FILING DATE: 2002-04-15
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 306
; LENGTH: 690
; TYPE: PRT
; ORGANISM: Homo Sapien
; US-10-123-155-306

```

Query Match	2.88; Score 33; DB 9; Length 690;
Best Local Similarity	9.8%; Pred. No. 1.7;
Matches	32; Conservative 100; Mismatches 195; Indels 0; Gaps 0;
QY	314 GGATAAAATCCGGCAGTCGCGCGGTCCATCAGACACATCCGCCACGGGTAAACCGT 373 : : : : : : : : : : : : : : : :
Db	401 GSADA.NSRA..GC.SWSNCNAH.SRWV.TG.SYRYHBMYS.A.A.M...R.A..NS 342 : : : : : : : : : : : : : : : :
QY	374 CCCTGTCCACATTTCTTGAATGATCATCAGGGATCCGCCGCTCTCACTGCGGATAACGGG 433 : : : : : : : : : : : : : : :
Db	341 SKSBS.BSS.BY.....SNHRN..M..NHSSS...C.SMRR..TBTENCNSHRTTMYM.. 282 : : : : : : : : : : : : : : :
QY	434 CAGCGCGGAGATGACGCTTCAGCGACTACCATACCAACGCTTCATTTCCGNAAGCAT 493 : : : : : : : : : : : : : : :
Db	281 .N.RHCHKND.D.DMKNRSH...BM.T.HASNAH..AAM..S.S.Y.AT..BADK..MA. 222 : : : : : : : : : : : : : : :
QY	494 GACCACCACTGGCATCCGGTAGACCGGTAAACGCTGGGAAAGGCACTGCCATTAA 553 : : : : : : : : : : : : : : :
Db	221 DAYYDNBS.MHWBB.AHY.B.NTBAMB..A.ASNS.AHWTS.ANNMRC..SSS.T....R 162 : : : : : : : : : : : : : : :
QY	554 CACATCTCCGCTCATTTCCCGAGTGTCTGTCTGTGACGACAGAGCTGCTTCGTATTCTTC 613 : : : : : : : : : : : : : : :
Db	161 A..HA.S.HABSNYRB....TB..M.SY.MA.AMN.NTT.GBNH.DGNTNBN..GBACHN 102 : : : : : : : : : : : : : : :
QY	614 ACGCCCGCGGCCACACGAGCCAGCG 640 : : : : : : : : : : : : : : :
Db	101 A..YH.NSSSY.C.BGKRRSC.ANAG 75 : : : : : : : : : : : : : : :

RESULT 29

```

RES001 23
US-09-918-995-5997
; Sequence 5997, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; TITLE OF INVENTION: FROM VARIOUS CDNA LIBRARIES
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5997
; LENGTH: 429
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-918-995-5997

```

Query Match 2.8%; Score 32.8; DB 9; Length 429;
Best Local Similarity 55.2%; Pred. NO. 1.4;
Matches 64; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

QY 358 CCACGGGTACAGCGTCCCTGTCACATCTTCTGAATGACATCAGGGATCCCGCCGCTT 417
DB 202 CCTGTGTAATCAGAGAGCCTGCCACATCTTGGAGTGTCCGAGGCGTTCGCCAGGATGT 261
QY 418 CACTGCGGATACGGGACGCGGAGACGTGACGCTTCACGAGTACCATACCAAAAC 473
DB 262 CATCAGCACCATGGCCAGCGCTTCGAGTGGCTTCAAAACATACCTCAGCAACC 317

RESULT 30

US-10-198-846-9370
; Sequence 9370, Application US/10198846
; Publication No. US2003009974A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinmann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; FILE REFERENCE: THERAPY OF BREAST CANCER
; CURRENT APPLICATION NUMBER: US/10/198,846
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306,220
; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 9370
; LENGTH: 617
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc_feature
; LOCATION: 15, 17, 27, 62, 165, 441, 544, 600
; OTHER INFORMATION: n = A,T,C or G
US-10-198-846-9370

Query Match 2.8%; Score 32.6; DB 9; Length 617;
Best Local Similarity 54.2%; Pred. No. 2.1;
Matches 65; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

QY 4 CAGGAGATGAAAAAGCCAAATATAAATGCCATCCAGCGCGCTCCAGTGAA 63
DB 370 CATGAGATTTAAAAACAAAAAAGCGCGCCAGCGCGTGGAGCTCCAA 429
QY 64 GTAGGCGCTTCTGTCGCGTATTAAATGCATTGACCGTCCCGCTATTAAACAATGTA 123
DB 430 TTCGCCCTATANTGAGTCTGATTACGCGGCTCAGTGGCGCTGTTTACACGCTGTGA 489

RESULT 31

US-10-184-644-414
; Sequence 414, Application US/10184644
; Publication No. US20030044930A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: ACIDS ENCODING THE SAME
; CURRENT APPLICATION NUMBER: US/10/184,644
; CURRENT FILING DATE: 2002-06-28

; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 414
; LENGTH: 811
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-184-644-414

Query Match 2.7%; Score 32.4; DB 9; Length 811;
Best Local Similarity 7.3%; Pred. No. 3;
Matches 25; Conservative 97; Mismatches 219; Indels 0; Gaps 0;

QY 756 GCCCAGCGGTGATATGCGAGGGGTAAAGCGTCCGATGCGTTCATATGCAATGCGGGCCAG 815
DB 423 VVNNLSYNKLSDSVFCPLPKSIQILDNNNQIOTVPKETIHLMLRELNLIAFNFLTDLP 482
QY 816 TCGAAACCCGGTGGATAACCGTTACCGGTGCTCGACACCTCCGCCATCAGATCGGCC 875
DB 483 GCSHFSRLSVLNIEMNFILSPSLDFVQSCQEVTLNAGRNPFRTCELKNFIOLEYSEV 542
QY 876 ATCATGGGTGATAGGCACAAATGAATCAACAGATATTAATTCAGGGAACAGTCTG 935
DB 543 MMVGSDSYTCYPLNLGRTRLDVHLHLSLNTALLIVIVIMLVGLAVAFCCCLHFD 602
QY 936 GTCCTACGGGTGATGATGAGTTTCTGCTGACAAATAGTAGAAGCGGTGACAGCATATCAGA 995
DB 603 LPWYRLMGQCTQTHRRVRKTTQELKRNVRHAFISYSRHSLSLVKNELIPNLEKEDGS 662
QY 996 CGGCTCAGTCCGCTATATACGTGCTGATGCGCACTATGCGAGATGACCATCAGGTGA 1055
DB 663 ILICLYESYFDGKSIENIVSFIKSYKSIFFLSNPNVONWCHVEFFFAHNLPHENS 722
QY 1056 AATTCGCCGATATCGTGAAGTCTGAGGATCGAAGGAAG 1096
DB 723 DHIIILLRPIPPYCIPTRYHKLKALLEKKAYLEWPKDRK 763

RESULT 32

US-10-184-634-414
; Sequence 414, Application US/10184634
; Publication No. US20030068684A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: ACIDS ENCODING THE SAME
; CURRENT APPLICATION NUMBER: US/10/184,634
; CURRENT FILING DATE: 2002-06-28
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 414
; LENGTH: 811
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-184-634-414

Query Match 2.7%; Score 32.4; DB 9; Length 811;
Best Local Similarity 7.3%; Pred. No. 3;
Matches 25; Conservative 97; Mismatches 219; Indels 0; Gaps 0;

QY 756 GCCCAGCGGTGATATGCGAGGGGTAAAGCGTCCGATGCGTTCATATGCAATGCGGGCCAG 815
DB 423 VVNNLSYNKLSDSVFCPLPKSIQILDNNNQIOTVPKETIHLMLRELNLIAFNFLTDLP 482

us-09-674-277-2.rnpb

RESULT 35
US-10-123-155-90
; Sequence 90, Application US/10123155
; Publication No. US20030068794A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J

```

; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330R1C30
; CURRENT APPLICATION NUMBER: US/10/123,155
; CURRENT FILING DATE: 2002-04-15
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 90
; LENGTH: 468
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-123-155-90

Query Match      2.7%; Score 32.2; DB 9; Length 468;
Best Local Similarity 7.0%; Pred. No. 2.4;
Matches 27; Conservative 117; Mismatches 241; Indels 0; Gaps 0;

QY 741 TTGCTCTGCTGATGAGCCAGCGGTGATATGACAGGGGTAAACGGTGCATGGCTTCATTA 800
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 58 YPENKPCQSNYSFVDNLNLLKATTEKEKIERQSISSPLDNKLNVEDVDSTKNRKLII 117
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 801 TGCATGCGGGCCAGTGCAGAACCCGGTGGAAATACCGTTACCGGTGCTCGACACCTTCC 860
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 118 DDYDSTKSLGDKHFKDDPDGLHQLDGTPLTAEDIVHKINARIYEENDRAVDFKIVSKLNL 177
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 861 GCCATCAGATCGCCATCATGGGTGAGATAGGACACAAATGAAATCACACAGATAATTC 920
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 178 LGLTESQAHVLEDAEVLQKLISKANNYEEDPNKPTSWENQAGKIKEKVTYPMIAIQ 237
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 921 AGGGAACAGCTTCTGGTCTTACGGGTGATGTAGTGTGTTTTTGTCTGCAATAGTAGAGCGG 980
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 238 DGLAKGENDETIVSNLTLTNGLERRTKTYSBDEPQFPNFPNFKLLSIDSEKEAKEKE 297
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 981 TGACAGCATACACAGCGCTCAGTCCCTGCTATATCTGTCATGCGCCACATATGCGAGATG 1040
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 298 TLITIMTKLYIDFVKMVKYGTISPVEGVSYLENDEMIALQTNKLNKLNATDNISKLPFA 357
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1041 ACCAGATCAGGTTTAAATCCCGCATATCCGTCGAAGTCTGAGGATGGAAGGAGGTGA 1100
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 358 PSEKSHETDSTKEAAKMEYGLSKDSTKDDNSNFGKTDPEKGTAEVLEAIRKNIE 417
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1101 AGGCTGTTCCTGAAAGGAATAAAG 1125
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 418 WLKKHDKKNGEDYDLKMRDFINK 442
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 36
US-09-738-626-451/c
; Sequence 451, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAOKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18

; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 451
; LENGTH: 1254
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-451

Query Match      2.7%; Score 32.2; DB 9; Length 1254;
Best Local Similarity 45.5%; Pred. No. 4.6;
Matches 115; Conservative 0; Mismatches 138; Indels 0; Gaps 0;

QY 270 TCTTCTCTGGCGCTGGTGCCTATCATCGGAGCGTTCCGGTCGGGATAAAAAATCGCGC 329
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1154 TCTTCGACGGCGTCTTTCACCCATTCTGATGCGCGTTTCGTATCATCTCCAGAGTGTGGCT 1095
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 330 ACTGCGCGGTCCATGCAGACACATCCCCACGCGTAAACAGCGTCCCTGTGCACATCTTTC 389
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1094 AAGGGTTCGGCCCGCATGCGGGGAGTGGCGTGCAGACGAATCCCGTCTCCCCCTTCC 1035
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 390 TGAATGACATCAGGATCCCGCCGCTCTCAGTTCGCGGATAACGGGACGCCGGAGACTGAC 449
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1034 GCGACTGCCATGGCAGGCGCCAAACCCGCGCGCAATGACCGGTGTGCGGTGGCTTGC 975
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 450 GCTTCAGCCAGTACCATACCAACGCTTCATTTCCGAAGCATGACACACACTGGCA 509
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 974 GCTTCATGCGCAGCAGTCCGAAGGACTCATTTAAACTTGGCAGCGGCACGATGTCGCGC 915
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 510 ATCCGCTAGACCG 522
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 914 GCCCGATACACG 902
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 37
US-09-917-800A-1734/c
; Sequence 1734, Application US/09917800A
; Patent No. US20020119462A1
; GENERAL INFORMATION:
; APPLICANT: Mendrick, Donna
; APPLICANT: Porter, Mark
; APPLICANT: Johnson, Kory
; APPLICANT: Castile, Arthur
; APPLICANT: Elashoff, Michael
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Molecular Toxicology Modeling
; FILE REFERENCE: 44921-5038-US
; CURRENT APPLICATION NUMBER: US/09/917,800A
; CURRENT FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: US 60/222,040
; PRIOR FILING DATE: 2000-07-31
; PRIOR APPLICATION NUMBER: US 60/222,880
; PRIOR FILING DATE: 2000-11-02
; PRIOR APPLICATION NUMBER: US 60/290,029
; PRIOR FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: US 60/290,645
; PRIOR FILING DATE: 2001-05-15
; PRIOR APPLICATION NUMBER: US 60/292,336
; PRIOR FILING DATE: 2001-05-22
; PRIOR APPLICATION NUMBER: US 60/295,798
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: US 60/297,457
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,884
; PRIOR FILING DATE: 2001-06-19
; PRIOR APPLICATION NUMBER: US 60/303,459
; PRIOR FILING DATE: 2001-07-09
; NUMBER OF SEQ ID NOS: 1740
```

```

; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 1734
; LENGTH: 1689
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020119462A1 NM_022403
US-09-917-800A-1734

Query Match          2.7%; Score 32.2; DB 10; Length 1689;
Best Local Similarity 52.6%; Pred. No. 5.6;
Matches 70; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

QY 1013 ATTACTGTCAGCCACTATGCGAGATGACGAGATCAGGTTTAATTTCCCGGTAATCCG 1072
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 351 ATCCCTGACATGGCCATTTTGAAAAATCTCAGCAACAGAAATCAAGTTCCCGAGAAATTTG 292
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1073 TCGAAGCTGAGGATGAAGAGGTGAAGCTGTTCTCTGAAAGGAATAAAGTGACATC 1132
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 291 TTTAACCAGAAGTTCATAAGCTTGGTGAGTTAATAAAGAGGTGCTGCTGGATTTT 232
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1133 ATGCCCTCTTTT 1145
    ||| ||| ||| |||
DB 231 ATCCCTTTGATT 219
    ||| ||| ||| |||

RESULT 38
US-09-738-626-1/c
; Sequence 1, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: Patent In ver. 3.0
; SEQ ID NO 1
; LENGTH: 3309400
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-1

Query Match          2.7%; Score 32.2; DB 9; Length 3309400;
Best Local Similarity 45.5%; Pred. No. 3.5e+02;
Matches 115; Conservative 0; Mismatches 136; Indels 0; Gaps 0;

QY 270 TCTTCTCGGCGTGGGTGCCATCGGGAAGGTTCCGTCGGGATATAAAAAATCGCGC 329
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 425031 TGTTCGAGCGGCTCTTCCACCATTTCTGATGCGGCTTTCGTATCGTCCAGAGTGTGCT 424972
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 330 AGTCGGCGGTCATGACAGACATCCCCACAGGGTAAACAGCGTCCCTGTGCATCTTC 389
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 424971 AAGCGTGGCCCGAGGATGCGGAGTGGCGGTCGACAAGCAATCCGCTCCCTTCC 424912
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 390 TGAATGACATCAGGATCCCGCGTCTCATCTGGGATAGCGGCGGACGACATGAC 449
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

DB 424911 GCGACTGGATGGCAGGCGCCCAACCCGCGCGGCGGCGGATGACCGGTGTGCCGTGCTTGC 424852
QY 450 GCTTCAGCCAGTACCATACCAAGCGCTTCATTTTCCGAAGCATGACCAACACACTGGCA 509
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 424851 GCCTCCATGGCGAGGAGTCCGGAAGGACTCATTTAAACCTTTGCGACGGCCACGATGTCGCC 424792
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 510 ATCCGGTAGACCG 522
    ||| ||| ||| |||
DB 424791 GCCCGATACACGG 424779
    ||| ||| ||| |||

RESULT 39
US-09-815-242-4609/c
; Sequence 4609, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4609
; LENGTH: 1311
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
US-09-815-242-4609

Query Match          2.7%; Score 32; DB 10; Length 1311;
Best Local Similarity 49.4%; Pred. No. 5.6;
Matches 83; Conservative 0; Mismatches 85; Indels 0; Gaps 0;

QY 633 AGCCAGCGAATGATTTCCCTTCCATCTTCCATGATACATAACGACGACATAATTC 692
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1166 ACCCAGGAATAGTGGCCACTTCATTTTAAATTTGTTTAAACACGACATAAGTTCA 1107
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 693 TGTCTTTTTCGGGACGTAGCATCCCGACCTGAACGATAAGCGGAACATTGTCTGCTGAT 752
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1106 TGCTCAGTACACCAATAAATCTTGAATTCATCAATTAATGGCATATATGTTTACCGAT 1047
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 753 GCAGCCAGGCGTGATGATGCGAGGGTAACGGTCCGATCGCTTCATTA 800
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1046 ACAGTAATGTTGGAAATAACGAAGAATGTTTCCAATCTTCATTA 999
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 40
US-09-815-242-8107/c
; Sequence 8107, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:

```

APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Cart, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
TITLE OF INVENTION: Prokaryotes
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIORITY APPLICATION NUMBER: 60/191,078
PRIORITY FILING DATE: 2000-03-21
PRIORITY APPLICATION NUMBER: 60/206,848
PRIORITY FILING DATE: 2000-05-23
PRIORITY APPLICATION NUMBER: 60/207,727
PRIORITY FILING DATE: 2000-05-26
PRIORITY APPLICATION NUMBER: 60/242,578
PRIORITY FILING DATE: 2000-10-23
PRIORITY APPLICATION NUMBER: 60/253,625
PRIORITY FILING DATE: 2000-11-27
PRIORITY APPLICATION NUMBER: 60/257,931
PRIORITY FILING DATE: 2000-12-22
PRIORITY APPLICATION NUMBER: 60/269,308
PRIORITY FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 8107
LENGTH: 1563
TYPE: DNA
ORGANISM: Staphylococcus aureus
FEATURE:
NAME/KEY: CDS
LOCATION: (1)...(1563)
US-09-815-242-8107

Query Match 2.7%; Score 32; DB 10; Length 1563;
Best Local Similarity 49.4%; Pred. No. 6.2;
Matches 83; Conservative 0; Mismatches 85; Indels 0; Gaps 0;
QY 633 AGCCAGGAATGATTTCCCTTCATCTTCACCTGATACAAATACAGCAGCATATCA 692
DB 1403 ACCAGGAATAGTTCCTTCACTTATTTTAAATTTGTTTAAACAGACATAAGTCA 1344
QY 693 TGTCCTTTTTCGGGACGTAGCATCCCACTGAACGATAAGCGGAACATTTCTGCTGAT 752
DB 1343 TGCTCACTACACCAATAATCTTGAATGATCAATAATGCGATATATGTTTACCGAT 1284
QY 753 GCAGCCAGCGGTGGATATGCGAGGGTAACGGTACGGTATGCTTCATTA 800
DB 1283 ACAGCTAATGTTTGGAAATAAGCAAGAATGTTTCCAAATCTTCATTA 1236

RESULT 41
US-10-184-644-402/c
; Sequence 402, Application US/10184644
; Publication No. US20030044930A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME

FILE REFERENCE: P3430RIC227
CURRENT APPLICATION NUMBER: US/10/184,644
CURRENT FILING DATE: 2002-06-28
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 612
SEQ ID NO 402
LENGTH: 653
TYPE: PRT
ORGANISM: Homo Sapien
US-10-184-644-402

Query Match 2.7%; Score 31.8; DB 9; Length 653;
Best Local Similarity 5.8%; Pred. No. 4.2;
Matches 34; Conservative 186; Mismatches 370; Indels 1; Gaps 1;
QY 290 CATCATCCGGAAGGTTCCGGTCCGGGATAAAATCGCGAGTCGCGCGTCCATCAGCA 349
DB 628 BD.SSH..B.C....HCS.CT...M.A.MM..RSBEC.HH.....SSSS..GYH.M..A 569
QY 350 CACATCCCCCAGGGTAACAGCGTCCTGTGTCATCTTCTGATGATCAGCAGGATCCC 409
DB 568 M.R..YH.MC..M.B..MDN..STSANT..B.RG.HMTN.H..NSSTYSDDYSSYS.SY 509
QY 410 GCCCGTCTCACTGCGGATAAGCGGACGCGGAGAGACTGACGCTTCAGCCAGTACCATACC 469
DB 508 NYMDMDYMDCCSSSSSSSSBS.HTSBSSSSSSMIYMM.YM..M..MY.M.TSA..MM 449
QY 470 AAACGCTTCATTTTCCGAAGCATGACCAACACACTCGCAATCCGGTAGACCGGTAAACGC 529
DB 448 AM.M..M.T.M.Y..SM.KRMD.M..TH.THM.A.H.T.TMRRS.TN..M.....CC 389
QY 530 TGGGAAAGGACCTCCCATTAACATCTCCGCTCATCTCCAGGTGTTCTGCTGCTG 589
DB 388 Y.B.A.G..RMVTNYDA..GN...T...H..TMN..SOMATR..TCYTB.T.BN.MH...TMNR. 329
QY 590 ACGCAGAGCTGCTTCTGATTTCTTCACGCGCGGCCCCACACGAGCGAGCGAAATGATTT 649
DB 328 KTHDYCBM.RHBC.MBGM.TWTS.MMY.TST.HH.S.NMS..CYK.S..NS.H.C.M 269
QY 650 CCCTTCCATCTTCAGCTGATACATAACAGCAGCATATAATCATGCTCTTTTCGGGACG 709
DB 268BBC.NB.C.SS.K-BN.RS..SNSN..BSYRR.....SS.BC.M.CS.D...SS 210
QY 710 TAGCATCCCCCAGCAAGGATAAGCGGAACATTTCTGCTGATGACGCCAGCGGTGAT 769
DB 209 SR.SBT.M.DRYH.HM.CTY..HCA.RRS..HCNDS.S.BHY..G.TA...DT.HYK.CS 150
QY 770 ATGAGGGGTAACGGTCCGATGCTTCAATGATGCAATGCGGGCCAGTCGAAACCCGGTGG 829
DB 149 C.G..BK..C.Y.SS.YC.B.HCY..HY....YVHKS...K.....K.TRDH.SH.N. 90
QY 830 AATAACCGTTACCGGTGCTTGACACCTTCGCGCATGATGCGCCATCAT 880
DB 89 .SAT.THSMWS.T...HTMMS..M....NHBYM.MIC....D....CYT..T 39

RESULT 42
US-10-184-634-402/c
; Sequence 402, Application US/10184634
; Publication No. US20030068684A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

FILE REFERENCE: CL1686 US NA
CURRENT APPLICATION NUMBER: US/09/891,641
CURRENT FILING DATE: 2001-06-26
NUMBER OF SEQ ID NOS: 81
SOFTWARE: Microsoft 97
SEQ ID NO 35
LENGTH: 1146
TYPE: DNA
ORGANISM: Bacillus subtilis
US-09-891-641-35

Query Match 2.7%; Score 31.8; DB 9; Length 1146;
Best Local Similarity 45.7%; Pred. No. 6;
Matches 111; Conservative 0; Mismatches 132; Indels 0; Gaps 0;

QY 272 TTCTCTGGCGTGGCTGCCATCATCCGGAAGCGTTCGGTCCGGATATAAAATCGCGCAG 331
DB 1050 TTTTGGGCTTCTGTCCTCCAGCTTTCGGCAGAGCTCCGGCTTATGGTAAAGCTGTTCAT 991
QY 332 TGGCGCGTCCATGCAGACACATCCCCACGGGTAAACAGCGTCCCTGTACATTTCTGTG 391
DB 990 CCGCGGGGCAAAAGCAGCACTGTCCACGATTTTGATCAGAAACCGGTTTCTCGGTGCGG 931
QY 392 AATGACATCAGGGATCCGCGCGCTCTCACTGGCGATACGGGCGACGCGGAGACTGACGC 451
DB 930 GATGATTTCCGATGCCGCGGATATCTGTGGCGATCGCGGTTGTTCTGCGCGCAATTC 871
QY 452 TTCAGCCAGTACCATAACAAACGGTTCATTTTCCGAAGGCGATGACACACACACTGGCAAT 511
DB 870 CTCAGCACATTCATACCGAGGCGCTTCTTAATGCTGGATGCGACAGATACATCCGCAAG 811
QY 512 CCG 514
DB 810 CTG 808

RESULT 45
US-09-759-130B-343/c
Sequence 343, Application US/09/759130B
Publication No. US2003002279A1
GENERAL INFORMATION:
APPLICANT: Millennium Pharmaceuticals, Inc.
APPLICANT: McCarthy, Sean A.
APPLICANT: Fraser, Christopher C
APPLICANT: Sharp, John D
APPLICANT: Barnes, Thomas S
APPLICANT: KIRST, Susan J
APPLICANT: Mackay, Charles R
APPLICANT: Myers, Paul S
APPLICANT: Leiby, Kevin R
APPLICANT: Wrighton, Nicolas
APPLICANT: Goodearl, Andrew
APPLICANT: Holtzman, Douglas A
TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING
TITLE OF INVENTION: PROGNOSTIC, DIAGNOSTIC, PREVENTIVE, THERAPEUTIC, AND OTHER
FILE REFERENCE: MPI00-5350NMIM
CURRENT APPLICATION NUMBER: US/09/759,130B
CURRENT FILING DATE: 2002-09-16
PRIOR APPLICATION NUMBER: US/09/479,249
PRIOR FILING DATE: 2000-01-07
PRIOR APPLICATION NUMBER: US/09/559,497
PRIOR FILING DATE: 2000-04-27
PRIOR APPLICATION NUMBER: US/09/578,063
PRIOR FILING DATE: 2000-05-24
PRIOR APPLICATION NUMBER: US/09/333,159
PRIOR FILING DATE: 1999-06-14
PRIOR APPLICATION NUMBER: US/09/596,194
PRIOR FILING DATE: 2000-07-14
PRIOR APPLICATION NUMBER: US/09/342,364
PRIOR FILING DATE: 1999-06-29
PRIOR APPLICATION NUMBER: US/09/608,452
PRIOR FILING DATE: 2000-06-30

PRIOR APPLICATION NUMBER: US 09/393,996
PRIOR FILING DATE: 1999-09-10
PRIOR APPLICATION NUMBER: US 09/602,871
PRIOR FILING DATE: 2000-06-23
PRIOR APPLICATION NUMBER: US 09/420,707
PRIOR FILING DATE: 1999-10-19
NUMBER OF SEQ ID NOS: 460
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 343
LENGTH: 3153
TYPE: DNA
ORGANISM: Mus sp.
FEATURE:
NAME/KEY: misc.feature
LOCATION: (1)..(3153)
OTHER INFORMATION: n - A,T,C or G
US-09-759-130B-343

Query Match 2.7%; Score 31.8; DB 9; Length 3153;
Best Local Similarity 51.0%; Pred. No. 11;
Matches 75; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

QY 396 ACATCAGGATCCCGCCGCTCTCACTGGCGATACGGGCGACCGGAGACTGACGGTTCA 455
DB 2000 ACTCCACCGTGGCTGGCACTGCGGTGGCAGCACTGGCTGGCCCTGCACTGAGGTGCT 1941
QY 456 GCCAGTACCATACCAACGCTTCATTTTCCGAAGGCGATGACCAACCACTGGCAATCCGG 515
DB 1940 GCCAGGACAGTTCTTCCAGACTTCTCTCTGAGGGGGTCTCCAGCTCCCTGGGACCCACA 1881
QY 516 TAGACGGGTACCGCTGGGAAAGGGCA 542
DB 1880 GGGGCCCCATGTAGCTGGAAGTAGGGAA 1854

RESULT 46
US-10-189-123-73/c
Sequence 73, Application US/10189123
Publication No. US20030082586A1
GENERAL INFORMATION:
APPLICANT: KIRST, Susan J.
APPLICANT: HOLTZMAN, Douglas A.
APPLICANT: FRASER, Christopher C.
APPLICANT: SHARP, John D. S.
APPLICANT: BARNES, Thomas S.
TITLE OF INVENTION: ANTIBODIES HAVING DIAGNOSTIC, PREVENTIVE, THERAPEUTIC, AND OTHER
FILE REFERENCE: 10147-11U3
CURRENT APPLICATION NUMBER: US/10/189,123
CURRENT FILING DATE: 2002-07-02
PRIOR APPLICATION NUMBER: US 09/596,194
PRIOR FILING DATE: 2000-06-16
PRIOR APPLICATION NUMBER: US 09/342,364
PRIOR FILING DATE: 1999-06-29
NUMBER OF SEQ ID NOS: 100
SOFTWARE: PatentIn version 3.1
SEQ ID NO 73
LENGTH: 3153
TYPE: DNA
ORGANISM: Mus sp.
FEATURE:
NAME/KEY: misc.feature
LOCATION: (3043)..(3043)
OTHER INFORMATION: unknown
FEATURE:
NAME/KEY: misc.feature
LOCATION: (3048)..(3048)
OTHER INFORMATION: unknown
US-10-189-123-73

Query Match 2.7%; Score 31.8; DB 9; Length 3153;
Best Local Similarity 51.0%; Pred. No. 11;
Matches 75; Conservative 0; Mismatches 72; Indels 0; Gaps 0;


```

Db      222 CC...CYV.TYBTMM...A.H.HSAM.S.SSS.SN....S.SBST.H.HSSTMTYMSBSKM 163
Qy      817 CGAACCCGGTGAATAACGGTTACC 842
Db      162 .T.AMYM.CSNHSSMHSHS.KYHS 137

RESULT 49
US-09-381-624A-5/c
; Sequence 5, Application US/09381624A
; Patent No. US20020068349A1
; GENERAL INFORMATION:
; APPLICANT: Horinouchi, Sueharu
; APPLICANT: Saitoh, Kohki
; APPLICANT: Takahashi, Eisaku
; TITLE OF INVENTION: GENE ENCODING TREHALOSE PHOSPHORYLASE, VECTOR CONTAINING THE GENE
; TITLE OF INVENTION: TRANSFORMANT TRANSFORMED BY THE GENE, AND METHOD FOR PRODUCING R
; TITLE OF INVENTION: TREHALOSE PHOSPHORYLASE WITH THE USE OF TRANSFORMANT
; FILE REFERENCE: FJIN-100
; CURRENT APPLICATION NUMBER: US/09/381.624A
; CURRENT FILING DATE: 1999-12-15
; PRIOR APPLICATION NUMBER: US 09/381.624
; PRIOR FILING DATE: 1999-12-15
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5
; LENGTH: 2199
; TYPE: DNA
; ORGANISM: Grifola frondosa
; NAME/KEY: promoter
; LOCATION: (1)..(3)
; NAME/KEY: terminator
; LOCATION: (2197)..(2199)
US-09-381-624A-5

Query Match      2.7%; Score 31.6; DB 10; Length 2199;
Best Local Similarity 46.4%; Pred. No. 11;
Matches 103; Conservative 0; Mismatches 119; Indels 0; Gaps 0;

Qy      290 CATCATCCGGAAGCGTTCCGGTCCGGGATAAAAAATCGCGCAGTGCGCCGGTCCATGCAGA 349
Db      1986 CATCGTGTACATACAGGTCCTCGTCCGTGTACAAAGTCGACATGTGCTGCGCAACTGCTGC 1927
Qy      350 CACATCCCCCAGGGTAACACGGTCCCTGTGCACATTCTTCGAATGACATCAGGGATCCC 409
Db      1926 GTTGTCGCCCGGCTCGCAGAGGTATCCGCTCTTTCCATGCTCAATCTGCAGCGGGATGCC 1867
Qy      410 GCCCGTCTCACTGGCGATAACGGCAGCCGCGAGACTGACGCTTCAGCCAGTACCATACC 469
Db      1866 GCCCGTGGCATGCGATGACGGGCTTGCCCTGCGTGAAGCCCTCTGAACCTTTACTTC 1807
Qy      470 AAACGCTTCATTTTCCGAAGGCATGACACCCACACACTGGCAAT 511
Db      1806 GAAGCCCTCACCGCTCGAGAGCTGTAGCGCGAACTTCGCATT 1765

RESULT 50
US-09-381-624A-4/c
; Sequence 4, Application US/09381624A
; Patent No. US20020068349A1
; GENERAL INFORMATION:
; APPLICANT: Horinouchi, Sueharu
; APPLICANT: Saitoh, Kohki
; APPLICANT: Takahashi, Eisaku
; TITLE OF INVENTION: GENE ENCODING TREHALOSE PHOSPHORYLASE, VECTOR CONTAINING THE GENE
; TITLE OF INVENTION: TRANSFORMANT TRANSFORMED BY THE GENE, AND METHOD FOR PRODUCING R
; TITLE OF INVENTION: TREHALOSE PHOSPHORYLASE WITH THE USE OF TRANSFORMANT
; FILE REFERENCE: FJIN-100
; CURRENT APPLICATION NUMBER: US/09/381.624A
; CURRENT FILING DATE: 1999-12-15
; PRIOR APPLICATION NUMBER: US 09/381.624
; PRIOR FILING DATE: 1999-12-15

```

```

; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 2817
; TYPE: DNA
; ORGANISM: Grifola frondosa
; NAME/KEY: promoter
; LOCATION: (1)..(65)
; NAME/KEY: Intron
; LOCATION: (177)..(239)
; NAME/KEY: Intron
; LOCATION: (486)..(547)
; NAME/KEY: Intron
; LOCATION: (1044)..(1160)
; NAME/KEY: Intron
; LOCATION: (1175)..(1228)
; NAME/KEY: Intron
; LOCATION: (1282)..(1350)
; NAME/KEY: Intron
; LOCATION: (1551)..(1618)
; NAME/KEY: Intron
; LOCATION: (1872)..(1928)
; NAME/KEY: Intron
; LOCATION: (2300)..(2356)
; NAME/KEY: terminator
; LOCATION: (2809)..(2811)
; NAME/KEY: terminator
; LOCATION: (2815)..(2817)
US-09-381-624A-4

Query Match      2.7%; Score 31.6; DB 10; Length 2817;
Best Local Similarity 46.4%; Pred. No. 13;
Matches 103; Conservative 0; Mismatches 119; Indels 0; Gaps 0;

Qy      290 CATCATCCGGAAGCGTTCCGGTCCGGGATAAAAAATCGCGCAGTGCGCCGGTCCATGCAGA 349
Db      2598 CATCGTGTACATACAGTCTCTCGTCCGTGTACAAAGTCGACATGTGCTGCGCAACTGCTGC 2539
Qy      350 CACATCCCCCAGGGTAACACGGTCCCTGTGCACATTCTTCTGAATGACATCAGGGATCCC 409
Db      2538 GTTGTCGCCCGGCTCGCAGAGGTATCCGCTCTTTCCATGCTCAATCTGCAGCGGGATGCC 2479
Qy      410 GCCCGTCTCACTGGCGATAACGGCAGCCGCGAGACTGACGCTTCAGCCAGTACCATACC 469
Db      2478 GCCCGTGGCATGCGATGACGGGCTTGCTGCGTGCAGAGGCTCTGAAACCTTTACTTC 2419
Qy      470 AAACGCTTCATTTTCCGAAGGCATGACACCCACACACTGGCAAT 511
Db      2418 GAAGCCCTCACCGCTCGAGAGCTGTAGCGCGAACTTCGCATT 2377

Search completed: June 19, 2003, 06:35:47
Job time : 311.567 secs

```